DOCUMENT RESUME

ED 267 624 FL 015 576

AUTHOR	Mace-Matluck, Betty J.; And Others
TITLE	Teaching Reading to Bilingual Children Study. Final Report.
INSTITUTION	Southwest Educational Development Lab., Austin, Tex.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	Nov 84
CONTRACT	400-83-0007
NOTE	1,102p.
AVAILABLE FROM	Southwest Educational Development Laboratory, 211 East Seventh Street, Austin, TX 78701.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF08/PC45 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Bilingual Education Programs; Bilingual Students; *Classroom Techniques; Elementary Education; English (Second Language); Federal Programs; Hispanic Americans; Language Proficiency; *Limited English Speaking; Longitudinal Studies; Native Language Instruction; *Reading Instruction; Reading Skills; Spanish Speaking; Speech Skills

ABSTRACT

This eight-volume final report of a federally-funded longitudinal study on the teaching of reading to bilingual children provides information on and insight into characteristics of a favorable learning environment for Spanish-speaking children, instructional sequences and events that promote successful and efficient learning of literacy skills, and the language and literacy outcomes of current schooling practices. The volumes deal with the following topics: (1) an introduction to and overview of the study; (2) a description of the study's design; (3) an outline of the analyses used to summarize patterns of growth in language skills, prereading skills, and reading comprehension, and to examine the link between instruction and achievement growth; (4) analyses of growth in oral language proficiency; (5) analyses of growth in reading skills; (6) data on the specific instructional techniques used in the classrooms studied; (7) a discussion of statistical techniques for integrating the findings on language, literacy, and instruction; and (8) the executive summary. (MSE)

Final Report

TEACHING READING TO BILINGUAL CHILDREN STUDY

Volume 1
Introduction

Betty J. Mace-Matluck
Wesley A. Hoover
Robert C. Calfee

Document BRS-84-R. 1-I

Preston C. Kronkosky
Executive Director

Southwest Educational Development Laboratory
211 East Seventh Street
Austin, Texas
(512) 476-6861

November 1984

Final Report

TEACHING READING TO BILINGUAL CHILDREN STUDY

Volume 1
Introduction

Betty J. Mace-Matluck, Wesley A. Hoover
Robert C. Calfee

Document BRS-84-R. 1-I

Preston C. Kronkosky, Executive Director
Southwest Educational Development Laboratory
Austin, Texas

November 1984

INTRODUCTION

> The ability to read well is basic to success in almost every aspect of the school curriculum. It is a prerequisite skill for nearly all jobs and is an important tool of lifelong learning (U.S. Commission on Civil Rights, 1971, P. 7).

Educators and policymakers alike have long recognized that the ability to read is essential for success in school, in work, and in life; yet many children from second-language backgrounds have trouble learning to read in schools today. The majority of these youngsters are from Spanish-language backgrounds and from low income families. Special programs designed to meet the needs of these children are provided in schools, but there is limited research evidence to guide the development, evaluation, and implementation of these programs.

Surveys of the general and school populations reveal an increase in the number of students whose language resources are not an ideal match to the language of the school. National statistics indicate that at present there are some 2.4 million American school children between the ages of five and fourteen who are from non-English language backgrounds and are limited in their English language abilities (O'Malley, 1982). These children present a special challenge to American educators, namely, to find effective ways of educating students who do not speak English sufficiently well to profit from instruction delivered in English. Limited English Proficient (LEP) students face a high risk of not completing school in this society (National Center for Educational Statistics, 1978; Waggoner, 1981; Steinberg, Blinde, & Chan, 1982). The school has frequently been the place where diverse groups of such children have been exposed to the language and common culture of the society. Through this exposure, many have found the means to become assimilated into the society; yet many have not. Language is only one of the many complex problems facing children from non-English backgrounds in the schools. Cultural backgrounds and social circumstances also influence the ways in which different groups respond to language differences they encounter in the classroom (Cohen, 1978). Nevertheless, drop-out rates among language minority students in our schools has always been relatively high. While it is not clear as to the extent to which language problems per se are responsible, what is clear is that language minority students, generally, do more poorly on standardized achievement tests and drop out earlier and in greater numbers than do their English monolingual classmates.

In an economy that has become increasingly technological, there are few jobs available for individuals who have less than a full education. Educational failure, whatever the cause, has become a major problem for the society; preventing it has become a major challenge for the schools. Language and related cultural differences have been identified as key factors in the educational failure of minority students, both by educators and by the courts (Leibowitz, 1982; Teitelbaum & Hiller, 1977; Cummins, 1981). An important question for educational practice and policy centers around the school's responsibilities in this situation.

Schooling for language minority students always has posed a particular challenge for American educators. Increasing numbers of such students, along with legislative mandates and greater attention being given to the educational plight of these students, have given rise to considerable debate about how U.S. school systems should respond to the cultural and linguistic diversity of their students. It has long been recognized that Limited English Proficient students are unable to participate fully in instruction delivered in English and that they need special assistance not only in (a) acquiring the necessary English skills to gain access to instruction but also in (b) making academic progress while those skills are being acquired. Special assistance, in the form of English-as-a-Second Language classes, has been provided over the years in some schools in an attempt to meet the first of these needs, but it was not until the passage of the Bilingual Education Act in 1968 that schools generally were encouraged to include instruction in the native language of the students to address the second of these (academic progress while acquiring the necessary English skills).

Bilingual education for language minority students proliferated in the early 1970s and expanded rapidly during the following decade. This expansion followed a landmark decision in 1974 (Lau vs. Nichols) in which the Supreme Court upheld the contention of a Chinese family that their child had been or was being denied access to equal educational opportunity because he was not sufficiently proficient in English to profit from instruction in English. Bilingual education, in which students are given instruction partially through their native language until they have attained sufficient proficiency in English to benefit from English medium instruction, was the principal remedy recommended by the Office for Civil Rights in response to the Supreme Court decision. From 1975 until very recently, school districts found to be out of compliance with the "Lau guidelines" could be denied access to federal education funds. During the early and middle 1970's a number of states also passed legislation mandating bilingual education and/or special language programs for limited English-speaking students.

With mandates from Congress and the courts that instruction in public schools take into consideration students' language and abilities, along with an increased awareness of the educational problems faced by children entering schools with limited English proficiency, educators have responded with instructional programs that are intended to provide equal access to the educational process. The goals of such programs are to concurrently develop English language proficiency while at the same time ensure progress in academic skills achievement. The best means by which to accomplish these goals has not been clearly established. The nature of the populations to be served, as well as local resources and educational philosophies, has given rise to a variety of organizational structures and instructional approaches for the delivery of this instruction (Mace-Matluck, Hoover, Domínguez, 1983). Although many individual programs have had considerable success in improving the academic performance of language minority students, it has not been demonstrated that these programs generally are reducing the educational failures of these students on the large scale that was envisioned. Thus, identifying more effective and practical means for increasing the academic success and building English proficiency for these students has been the focus of a number of studies funded over the past several years by the National Institute of Education. One of these is the Teaching Reading to Bilingual Children study described below.

Description of the Study

In June 1978 the National Institute of Education funded the Southwest Educational Development Laboratory (SEDL) to conduct a longitudinal study on the teaching of reading to bilingual children. The purpose of the study was to provide information that could result in greater insights into what constitutes a favorable learning environment for children from Spanish-language backgrounds, what instructional sequences and events promote successful and efficient learning of literacy skills, and what the language and literacy outcomes of current schooling practices are for a large sample of these youngsters.

Growth in reading comes about for most youngsters through formal classroom instruction. Understanding the development of reading, and knowledge of the critical variables that determine success or failure, depends on a careful examination of the instructional program -- not just the label over the classroom door, but the program as actually implemented by the classroom teacher.

Educators have raised several issues about the most effective way to help bilingual children become proficient readers of English. These include (a) valid assessment of the student's ability in the languages of the home and of the school, (b) the optimal balance of formal instruction in both languages, (c) the most effective transfer of skills from one language to the other, and (d) bilingual support within the classroom environment. A major thesis of the Teaching Reading to Bilingual Children study is that addressing these issues (and other requires a comprehensive and ecologically-valid investigation of the linkage between the child's language and the language of instruction.

It is well documented that, in general, children from Spanish-language backgrounds, for whatever reason, often encounter difficulty in our nation's schools; they do more poorly on standardized tests than does the general school population, and their dropout rate is high. Moreover, Hispanics make up the largest and fastest growing school-age population today. The demographics for some states show that over the next decade they may constitute as much as a third to a half of the population. In the six-state region served by the Southwest Educational Development Laboratory (Arkansas, Louisiana, Mississippi, Oklahoma, New Mexico, and Texas), Hispanics, whose backgrounds are tied to Mexico (Mexicans and Mexican Americans), have long been the largest single minority group. Two of the states, Texas and New Mexico, rank among the ten states most active in bilingual education (third and eighth respectively) as indicated by the level of ESEA Title VII funds allocated for such programs. In the state of Texas at present approximately one third of the school children are from Hispanic backgrounds (approaching one million). They are found in virtually every school district in the state. Many of the school districts in the southern portion of the state serve school populations of which 75% to 99% of the children are from Spanish-speaking backgrounds and, on entry into school, are often limited in their ability to speak English and to profit from instruction in that language. This population is not restricted to the border areas, however. Large urban centers in the state report as much as 20% of their school population from Hispanic backgrounds, with a concentration of some 80% to 90% in certain of their schools. Similarly, the population in New Mexico's public schools is heavily Hispanic. Approximately 61% of the children in grades one through three are from non-English-language backgrounds and are provided special language assistance programs. Of the 36,000 students in these programs, the large majority is Hispanic. In the other four states served by SEDL, clusters of Hispanic school children are identifiable, but in considerably smaller numbers than in Texas and New Mexico.

The study, conducted during the years of 1978 through 1984, focused on Spanish-speaking children from low income families in Texas. It is a comprehensive longitudinal investigation of the development of reading skills from kindergarten through fourth grade for a representative sample of more than 350 children from bilingual backgrounds, and for smaller samples of children who, on entry into school, were monolingual in English or Spanish. In this "natural variation" study, teaching and learning were carefully documented in field settings at the several sites.

The goals of the study were to (a) describe variations in both English and Spanish language ability of students living in bilingual communities, (b) document prevailing practices in reading instruction for bilingual students, and (c) investigate the relations between the instructional program and student achievement for students with differing entry profiles.

Design of the Study

Historical Perspective

In 1977, when the Teaching Reading to Bilingual Children study was being designed, a number of forces, in addition to the ones discussed above, were visible on the educational scene. Reading educators were debating the value of "phonics" in teaching children to read; debate ensued between the theorists who espoused the information process models of reading and those who espoused the analysis-by-synthesis models. Further complicating the matter for teachers responsible for teaching reading to Spanish dominant children was the prevailing notion that it is easier to learn to read initially in a language in which the sound-symbol correspondences are relatively regular, as is the case with Spanish. It was further claimed that reading is a single process and, that once having gained skill in reading Spanish, for example, these skills could be transferred by the child to reading English text. Little guidance was given, however, in what exactly it is that is transferable, just how transfer takes place, or what instructional practices facilitate transfer of learning.

Similarly, cognitive style research had been expanded to include the educational process. "Evidence" was being put forth that Mexican Americans children scored higher on the traits of field dependence and impulsivity than did the general population. It was therefore claimed that, to ensure equal access to education, a special curriculum accompanied by specified instructional techniques was required for these children.

Issues such as the above were the topic of discussion of a group of bilingual educators and researchers convened at SEDL in the spring of 1977. From that discussion came a research agenda; high priority among the topics identified was a longitudinal study that would investigate a number of questions related to the issues surrounding the teaching of reading to bilingual children.

In the original proposal submitted to the National Institute of Education, issues related to cognitive style and reading methodology were the primary focus. However, it became clear in the early years of the study that (a) the construct of "cognitive style" was not clearly defined, (b) Hispanic children in the study exhibited the full range of scores on selected measures of cognitive style, thus, the tendency toward one trait or another did not hold among the children in the sample, (c) no evidence was found that scores on the cognitive style measures were predictors of or related to student achievement in reading, and (d) reading practices prevalent in the schools could not be characterized as one reading "methodology " or another, but rather a combination of several. The SEDL research staff continued to carry out the original design of the study but began to turn their attention to factors which appeared to be most relevant to the purpose of the study (i.e., language characteristics of the students, reading-related knowledge and skills on entry into school, growth patterns in language and reading, the nature of the instruction and of the instructional program, and the relations between the instructional program and student achievement for students with differing entry profiles).

Methodology

To achieve the objectives of the study, considerable attention was given to the selection of schools, teachers and students, to the instruments for assessing language and reading achievement, and to the methods for evaluating the classroom instruction. Each of these topics is discussed briefly below.

Schools, classes and teachers. Twenty schools and 200 teachers from six school districts participated in the study. Included were variations in the nature of the reading program (a range from phonics-oriented to meaning-based), classroom organization (some self-contained, others team-taught), and grade structure (the range of grades in the individual school and the extent of cross-grading both vary). The schools differed in size, SES, urbanicity, locale, and makeup of the student body (from medium to high concentration of bilingual students).

Student cohorts. The study was undertaken in four cohorts or "waves" of students. Three of the cohorts consisted entirely, or in large part, of bilingual students. The first was small (N=40) and of limited generality; the second was somewhat larger (N=80) and covered a slightly broader array of contexts. The third cohort which was both larger (N=200) and broader in its generality, incorporated a number of procedural improvements based on previous experience in the study and included a monolingual English-speaking sample. The fourth cohort consisted of a relatively small sample (N=60) of monolingual Spanish-speaking students.

All of the bilingual sites were from the state of Texas, as were the monolingual English-speaking students. The monolingual Spanish-speaking students were from one site in Northern Mexico. The original design of the study called for each student to be assessed and observed from entry to kindergarten through exit from third grade. By covering the full range of the primary years, we would be able to examine the transition from "learning to read" through "reading to learn." For students in programs where the initial stages of reading were in Spanish, we also considered it important to determine the transition to competence in English reading.

The original design was in fact implemented for the first two cohorts; some of the students were tracked from first through fourth grade, but most followed the intended design. Due to limited funding in the later stages of the study, the last two cohorts could not be followed for the full four years that were originally intended. The bilingual and monolingual English samples from the Texas sites were observed from kindergarten through second grade, and the monolingual Spanish samples from the site in Northern Mexico were observed from first through third grade (the program did not provide a kindergarten).

The monolingual samples were incorporated in the design to aid in validating the instruments for student assessment. Both the English and Spanish cohorts are small and not selected to be fully representative of monolingual populations. Data from these samples will be presented in Volume 3, as part of the discussion on the adequacy of the instruments for measuring growth. The study was designed to study the course of reading in bilingual students, not as a basis for comparing these students with monolingual youngsters. Accordingly, comparisons between the various samples will not be made in this report, nor do we recommend that others attempt such comparisons.

Language assessment. Several types of data were collected for each student on English and Spanish proficiency. Each year, early in the Fall and again in the Winter and Spring, teachers rate their students' language skills. Oral language proficiency tests were administered in the Fall of each year. Finally, audiotaped speech samples were obtained monthly on a rotating schedule in three settings: in the classroom, on the playground, and in the home.

Reading Assessment. Several instruments were used to measure reading achievement. Standardized test scores (mostly English) were collected yearly. More detailed information was obtained from a battery of individually-administered "performance-based tests" in both English and Spanish. In kindergarten, the Stanford Foundation Skills Test was employed to measure the child's pre-reading skills. From the end of first grade on, the Interactive Reading Assessment System was administered during the Spring of each school year. This instrument provides independent measures of the student's skills in decoding, word meaning, fluency in oral reading, and comprehension. Finally, informal reading inventories were administered throughout the school year.

Classroom observations and teacher interviews. Project staff conducted monthly observations of the reading instruction in each classroom and interviewed the teachers quarterly about their instructional plans. The observation instrument documented staffing patterns, grouping and organization, time allocation, the language of instruction, the character of instruction, the materials and procedures used, and the response of the students. The interviews focused on the teacher's general instructional objectives, as well as the objectives for individual target students. Taken together, these two instruments yielded a rich characterization of the classroom environment for the target students.

Student entry variables, classroom factors, and reading achievement. The primary goals of the analyses were to identify the general relationships that characterize variation in these factors and to look for underlying regularities that are associated with success and failure, both in the early stage of reading instruction and in the year-to-year variations.

Documents

This report is one of a series of eight documents contained in the Final Report submitted to the National Institute of Education. A complete list of these documents is provided on the inside of the cover of this report.

The study was a collaborative effort among a number of individuals and institutions. All members of the research team contributed to the thinking, planning, and writing of this series of documents, however, the individual whose name appears first in the list of authors was responsible for preparing the particular document.

References

Cohen, A. D. (1978). Researching the linguistic outcomes of bilingual programs. The Bilingual Review, 9, (2), 97-108.

Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California State Department of Education, Schooling and language minority students: A theoretical framework. Los Angeles, CA: National Dissemination and Assessment Center.

Leibowitz, A. H. (1982). Federal recognition of the rights of minority language groups. Rosslyn, VA: National Clearinghouse for Bilingual Education.

Mace-Matluck, B. J., Hoover, W. A., & Domínguez, D. (1983). Language and literacy learning in bilingual instruction: A case study of practices and outcomes (Contract No. 400-80-0043). Washington, DC: National Institute of Education.

National Center for Educational Statistics. (1978). The education of language-minority persons in the United States, Spring 1976. NCES Bulletin, 78-B-4, July.

O'Malley, J. M. (1982). Children's English language and services study: Educational needs assessment for language minority children with limited English proficiency. Rosslyn, VA: National Clearinghouse for Bilingual Education.

Steinberg, L., Blinde, P. L., & Chan, K. S. (1982). Dropping out among language-minority youth: A review of the literature. NCBR Working Papers, No. 81-3W. Los Alamitos, CA: National Center for Bilingual Research.

Teitelbaum, H., & Hiller, R. J. (1977). The legal perspective. Bilingual education: Current perspectives Vol. 3. Arlington, VA: Center for Applied Linguistics.

United States Commission on Civil Rights. (1971). The unfinished education: Outcomes for minorities in the five southwestern states, Mexican American Educational Series, Report II. Washington, DC: U. S. Government Printing Office.

Waggoner, D. (1981). Educational attainment of language minorities in the United States. NABE Journal, 6, 41-53.

Final Report

TEACHING READING TO BILINGUAL CHILDREN STUDY

Volume 2
Design of the Study

Betty J. Mace-Matluck
Wesley A. Hoover
Robert C. Calfee

Document BRS-84-R. 1-II

Preston C. Kronkosky
Executive Director

Southwest Educational Development Laboratory
211 East Seventh Street
Austin, Texas
(512) 476-6861

November 1984

'inal Report

TEACHING READING TO BILINGUAL CHILDREN STUDY

Volume 2
Design of the Study

Betty J. Mace-Matluck, Wesley A. Hoover
Robert C. Calfee

Document BRS-84-R. 1-II

Preston C. Kronkosky, Executive Director
Southwest Educational Development Laboratory
Austin, Texas

November 1984

There were many individuals and institutions who contributed to this research effort. We wish to express our sincere gratitude to the parents, students, and school personnel who provided the necessary data from which this study is derived.

In addition, several other individuals made valuable contributions to the study, for which we are indebted: Robert C. Calfee, Sylvia C. Peña, and Blanca de Alvarez.

And finally, we wish to thank the local data collectors at the school sites, many of whom remained with the study throughout its duration: Ramiro Barrera, Beatrice Cantú. Irene Cavazos, Carolyn Cruz, María de Obregón, Gloria de Torres, Gigi Galván, Olga Hernández, Irene Méndez, Guadalupe Treviño, Rosalinda Villanpando, and Gloria Villarreal. Their patience, dedication, and hard work helped make this study a reality.

Betty J. Mace-Matluck
Wesley A. Hoover

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

16

PREFACE

In June 1978 the National Institute of Education (NIE) funded the Southwest Educational Development Laboratory (SEDL) to conduct a longitudinal study on the Teaching of Reading to Bilingual Children. Educators and policymakers alike have long recognized that the ability to read is essential for success in school, in work, and in life; yet many children from second-language backgrounds have trouble learning to read in schools today. The majority of these youngsters are from Spanish-language backgrounds and from low income families. Special programs designed to meet the needs of these children are provided in schools, but there is limited research evidence to guide the development, evaluation, and implementation of these programs. This study is intended to provide information that will result in greater insights into what constitutes a favorable learning environment for children from Spanish-language backgrounds, what instructional sequences and events promote successful and efficient learning of literacy skills, and what the language and literacy outcomes of current schooling practices are for a large sample of these youngsters.

The study was conducted during the years of 1978 through 1984. It is a comprehensive longitudinal investigation of the development of reading skills from kindergarten through fourth grade for a representative sample of more than 350 children from bilingual backgrounds, and for smaller samples of children who, on entry into school, were monolingual in English or Spanish. In this "natural variation" study, teaching and learning were carefully documented in field settings at the several sites.

The goals of the study were to (a) describe variations in both English and Spanish language ability of students living in bilingual communities, (b) document prevailing practices in reading instruction for bilingual students, and c) investigate the relations between the instructional program and student achievement for students with differing entry profiles.

Description of the Study

Surveys of the general and school populations reveal an increase in the number of students whose language resources are not an ideal match to the language of the school. An important question for educational practice and policy centers around the school's responsibilities in this situation. Bilingual programs, English-as-a-Second-Language classes, classroom aides, and "sink-or-swim" approaches can all be found in practice today. From limited evidence now available, none of these techniques has emerged as the one best system.

Hispanics make up the largest and fastest growing school-age population today. The demographics for some states show that over the next decade they may constitute as much as a third to a half of the population. In the state of Texas at present approximately one third of the school children are from Hispanic backgrounds (approaching one

million). They are found in virtually ever school district in the state. Many of the school districts in the southern portion of the state serve school populations of which 75% to 99% of the children are from Spanish-speaking backgrounds and, on entry into school, are often limited in their ability to speak English and to profit from instruction in that language. This population is not restricted to the border areas, however. Large urban centers in the state report as much as 20% of their school population from Hispanic backgrounds, with a concentration of some 80% to 90% in certain of their schools.

It is well documented that, in general, children from Spanish-speaking backgrounds, for whatever reason, often encounter difficulty in our nation's schools; they do more poorly on standardized tests than does the general school population, and their dropout rate is high. Bilingual education, in which students are given instruction partially through the home language until they have attained sufficient proficiency in English to benefit from English-medium instruction, has been the principal approach recommended by the Office for Civil Rights to ensure access to equal educational opportunity for these children. Although many individual programs have had considerable success in improving the academic performance of language-minority students, it has not been demonstrated that these programs generally are reducing inequality of educational opportunity on the large scale that was envisioned.

Growth in reading comes about for most youngsters through formal classroom instruction. Understanding the development of reading, and knowledge of the critical variables that determine success or failure, depends on a careful examination of the instructional program -- not just the label over the classroom door, but the program as actually implemented by the classroom teacher.

Educators have raised several issues about the most effective way to help bilingual children become proficient readers of English. These include (a) valid assessment of the student's ability in the languages of the home and of the school, (b) the optimal balance of formal instruction in both languages, (c) the most effective transfer from one language to the other, and (d) bilingual support within the classroom environment. A major thesis of the Teaching Reading to Bilingual Children study is that addressing these issues (and others) requires a comprehensive and ecologically-valid investigation of the linkage between the child's language and the language of instruction.

Design of the Study

To achieve the objectives of the study, considerable attention was given to the selection of schools, teachers and students, to the instruments for assessing language and reading achievement, and to the methods for evaluating the classroom instruction. Each of these topics is discussed briefly below.

Schools, Classes and Teachers

Twenty schools and 200 teachers from six school districts participated in the study. Included are variations in the nature of the reading program (a range from phonics-oriented to meaning-based), classroom organization (some self-contained, others team-taught), and grade structure (the range of grades in the individual school and the extent of cross-grading both vary). The schools differed in size, SES, urbanicity, locale, and makeup of the student body (from medium to high concentration of bilingual students).

Student Cohorts

The study was undertaken in four cohorts or "waves" of students. Three of the cohorts consisted entirely, or in large part, of bilingual students. The first cohort was small (N=40) and of limited generality; the second was somewhat larger (N=60) and covered a slightly broader array of contexts. The third cohort which was both larger (N=200) and broader in its generality, incorporated a number of procedural improvements based on previous experience in the study and included a monolingual English-speaking sample. The fourth cohort consisted of a relatively small sample (N=60) of monolingual Spanish-speaking students.

All of the bilingual sites were from the state of Texas, as were the monolingual English-speaking students. The monolingual Spanish-speaking students were from one site in Northern Mexico.

The original design of the study called for each student to be assessed and observed from entry to kindergarten through exit from third grade. By covering the full range of the primary years, we would be able to examine the transition from "learning to read" through "reading to learn." For students in programs where the initial stages of reading were in Spanish, we also considered it important to determine the transition to competence in English reading.

The original design was in fact implemented for the first two cohorts; some of the students were tracked from first through fourth grade, but most followed the intended design. Due to limited funding in the later stages of the study the last two cohorts could not be followed for the full four years that were originally intended. The bilingual and monolingual English samples from the Texas sites were observed from kindergarten through second grade, and the monolingual Spanish samples from the site in Northern Mexico were observed from first through third grade (the program did not provide a kindergarten).

The monolingual samples were incorporated in the design to aid in validating the instruments for student assessment. Both the English and Spanish cohorts are small and not selected to be fully representative of monolingual populations. Data from these samples will be presented in Volume 3, as part of the discussion on the adequacy of the instruments for measuring growth. The study was designed to study the course of reading in bilingual students, not as a basis for comparing these students with monolingual youngsters. Accordingly, comparisons

between the various samples will not be made in this report, nor do we recommend that others attempt such comparisons.

Language Assessment

Several types of data were collected for each student on English and Spanish proficiency. Each year, early in the Fall and again in the Winter and Spring, teachers rated their students' language skills. Oral language proficiency tests were administered in the Fall of each year. Finally, audiotaped speech samples were obtained monthly on a rotating schedule in three settings: in the classroom, on the playground, and in the home.

Reading Assessment

Several instruments were used to measure reading achievement. Standardized test scores (mostly English) were collected yearly. More detailed information was obtained from a battery of individually-administered "performance based tests" in both English and Spanish. In kindergarten, the Stanford Foundation Skills Test was employed to measure the child's pre-reading skills. From the end of first grade on, the Interactive Reading Assessment System was administered during the Spring of each school year. This instrument provides independent measures of the student's skills in decoding, word meaning, fluency in oral reading, and comprehension. Finally, informal reading inventories were administered throughout the school year.

Classroom Observations and Teacher Interviews

Project staff conducted monthly observations of the reading instruction in each classroom and interviewed the teachers quarterly about their instructional plans. The observation instrument documented staffing patterns, grouping and organization, time allocation, the language of instruction, the character of instruction, the materials and procedures used, and the response of the students. The interviews focused on the teacher's general instructional objectives, as well as the objectives for individual target students. Taken together, these two instruments yield a rich characterization of the classroom environment for the target students.

Student Entry Variables, Classroom Factors, and Reading Achievement

The primary goals of the analyses were to identify the general relationships that characterize variation in these factors and to look for underlying regularities that are associated with success and failure, both in the early stage of reading instruction and in the year-to-year variations.

Documents

This report is one of a series of eight documents contained in the Final Report submitted to the National Institute of Education. A com-

plete list of these documents is provided on the inside of the cover of this report.

The study was a collaborative effort among a number of individuals and institutions. All members of the research team contributed to the thinking, planning, and writing of this series of documents, however, the individual whose name appears first in the list of authors was responsible for preparing the particular document.

Betty J. Mace-Matluck
Wesley A. Hoover
Co-Principal Investigators

Austin, Texas
November 30, 1984

Introduction

The primary purpose of the study was to gain an understanding of the processes by which children from Hispanic backgrounds (more specifically, children likely to have acquired Spanish during their preschool years, and to be limited in their English skills on entry into school) learn to read while enrolled in schools which have adopted a variety of strategies to aid the children in becoming fluent readers of English.

The major goals of the study were to (a) describe variations in both English and Spanish language abilities of students living in bilingual communities; (b) [illegible]nt prevailing practices in classroom instruction for bilingual students; and (c) investigate the relations between the instructional programs and student achievement for students with differing entry profiles.

Questions and Hypotheses

The study, as originally proposed, was aimed toward answering three generic questions:

1. What are the effects of learner characteristics and reading instruction on the acquisition of literacy by bilingual students?

2. What is the extent of transferability of literacy in Spanish to literacy in English during the acquisition of literacy by bilingual students?

3. What is the effect of variations in instructional methodology (e.g., early emphasis on decoding versus comprehension) on the acquisition of literacy by bilingual children?

These broad-based questions, together with the theoretical rationale discussed below, provided the foundation for the design of the study. However, more specific hypotheses were required for planning specific contrasts in the data structure. These included the following:

Hypothesis 1. Given constant levels of pre-reading skill and instruction, students who enter school with higher scores in language proficiency will show greater rates of gain in the early stages of reading achievement, especially in the language-related areas of achievement such as vocabulary and listening comprehension.

Hypothesis 2. Given constant levels of school-entry language proficiency and instruction, students with higher pre-reading scores will show greater rates of gains in the early stages of reading achievement.

Hypothesis 3. Given constant levels of pre-reading skill and instruction, children who are fluent in both English and Spanish will show higher rates of gain in the acquisition of literacy over the long run.

Hypothesis 4. The rate of growth in language proficiency will parallel the rate of growth in reading achievement.

Hypothesis 5. For constant levels of the precursor (entry-level) student characteristics, well-managed, text-oriented instruction will result in greater immediate gains in reading achievement; the domains of reading in which the gains are greatest will correspond to the instructional focus.

Hypothesis 6. For constant levels of the precursor student characteristics, instructional emphasis on decoding principles in the early primary grades will result in higher rates of gain in reading achievement, especially (if obviously) in the decoding component, but perhaps in other areas as well.

Hypothesis 7. For constant levels of language proficiency in both English and Spanish, higher levels of reading achievement in Spanish will be correlated with more rapid rates of gain in English reading achievement.

The above hypotheses, which are related in various ways to the three basic questions, reflect analyses of the research literature on reading acquisition by bilingual and monolingual children, as well as input from several consultants, both researchers and practitioners.

Theoretical Rationale

Bilingual reading instruction, in addition to an emphasis on bilinguality, also entails a focus on reading and on instruction. The theoretical model relied upon for these latter two elements is the separable-process model proposed by Calfee (1977) as a technique for decomposing the complexities of reading and formal instruction. Two additional important concepts underlying the design of the study were the contrast between formal and natural language and teaching for transfer. A discussion of each of these follows.

Separable-Process Model of Reading and Reading Instruction

The instruments for assessing reading and observing instruction in the study were designed on the basis of principles of cognitive psychology. The last quarter-century has seen a revolution in psychology, with the change from emphasis on the external facets of behavior toward an exploration of the mental processes that underlie performance. Both teachers and students are "thinkers, " and one can represent the process of education as a matter of changing minds. Accordingly, it seemed appropriate to ground the present study in the findings from research on human information-processing. The summary that follows is necessarily abbreviated; for a more complete discussion, see Calfee (1981) and the references provided therein.

While it is easy to be awed by the apparent complexities of human thought, and while many studies of cognition seem quite complicated, the major theme that is now springing from cognitive research is one of simplicity:

> A few basic characteristics of the human information-processing system shape its problem-solving efforts. Apart from the sensory organs, the system operates almost entirely serially, one process at a time, rather than in parallel fashion. This seriality is reflected in the narrowness of its momentary focus of attention... Inputs and outputs... are held in a small short-term memory with a capacity of only a few (between, say, four and seven) familiar symbols or chunks. The system has access to an essentially unlimited memory.... (Simon, 1978, p. 273).

Simon's last point merits further comment. The mind, as an organ for storing and organizing experience, has a capability that is for practical purposes unlimited. Storage is influenced by the well-known variables of frequency and similarity (the mind operates in a content-addressable manner, unlike digital computers which operate according to a location-addressable principle). Organization of information in human memory can occur naturally, through the aggregation of common experiences, but it can also occur as the result of formal instruction. In either event, we are extremely limited in the number of distinctive experiences (whether present or remembered) that we can think about at any given moment. The limited-capacity feature of human information-processing has a number of consequences for instruction, and for research on instruction. The mind cannot encompass complicated, multifaceted phenomena without imposing some simple structure on them. Sternberg (1967) and Calfee (1977) have used the term independent processes to refer to the separability of elements in human thought; Simon (1981) talks about the decomposability of the components in any complex system, including the mind. In all of these instances, the major conclusion is the need to break a complicated "whole" into a small number of relatively separable parts, such that the interrelations between the parts is relatively simple, even though the interactions within a part may be relatively complex. A decomposed system is "comprehensible" -- the mind can grasp it, even though the mind has a limited capacity. Over the course of human history, a number of complexities have yielded to the decomposition principle with such success that they have become important parts of the school curriculum (e.g., Newton's laws of motion, the biological taxonomies, and the theory of the universe, to mention a few examples). In general, any time something can be made simple, the reduction merits serious consideration.

Reading and instruction pose quite a challenge, given the preceding remarks. There is a tendency for reading experts to emphasize the complexity of reading, and for researchers of classroom instruction to stress the multivariate character of teaching. While admitting to the surface complexity of both reading and reading instruction, the task taken in the design of the present study was that it was essential to plan the design of the study around a simple representation of these domains. For reading, the representation chosen was the independent-process model of reading proposed by Calfee (1977), according to which the basic components underlying the performance of a competent reader were four in number -- decoding, word meaning, propositional comprehension, and text comprehension. The substance of these components is spelled out in the reference listed; briefly, the components are sufficiently distinct to allow the construction of a reading test and the design of a classroom observation schedule in which each of the components has independent status.

For instruction, the structural model of teaching described by Calfee and Shefelbine (1981) was taken as a guide. In this model, the competent teacher is described as possessing separable domains of knowledge in the areas of knowledge of the curriculum, concepts of learning, analysis of instructional materials, assessment and diagnostic techniques, long-term management, and interactional principles. Each of these components is represented in the design of the classroom observation system.

While the concept of separable processes in reading and reading instruction has played an important role in the design of the instrument package for the study, and has served as an important guide in the analysis of the data structure, this concepts needs to be placed in proper perspective. First, the concept (or theoretical model, if you will) has the status of a hypothesis. The data of the study would either support the hypothesis (the findings would be reasonably coherent and interpretable), or not. Second, the model has a semi-normative character. The competent reader and the competent teacher, as the model has been formulated, operates according to the independent-process principle. We expected that some readers and some teachers would not reflect the separability of elements called for in the model -- we would expect such readers to perform poorly, and such teachers to be less effective in training their students to read.

Formal and Natural Language

In addition to the concept of separable processes in reading, a second important theme underlying the design of the study has been the contrast between formal and natural language. This contrast, which some scholars (Goody & Watt, 1963; Olson, 1977, 1980) link to the distinction between writing and speaking, has important consequences for education. Indeed, the argument can be made that an important characteristic of the competent adult in modern society is the ability to deal with the various manifestations of language in a formal way -- to treat language as a tool. The tendency is to equate formal language with the medium of print, since for some time reading and writing have constituted the main vehicle for teaching the techniques of formal language. Formal competence, however, permeates all manner of language usage -- speaking, discussing, even listening (Calfee, 1982, Calfee & Sutter, 1982; Cummins, 1980; Heath, in press). It is, moreover, not just a different style of language, but a different way of thinking.

The major distinctions between the natural and formal styles are summarized in Figure 1. These distinctions, which should be viewed as endpoints on a continuum, provide a useful framework for evaluation of the degree of "schoolness" in any particular linguistic situation. Most of the contrasts should be self-evident: additional detail is provided in the references listed above. The first two distinctions, however, merit further comment.

In formal language, little is left to chance. Misinterpretation is kept to a minimum, whenever feasible, by a high degree of explicitness. People have a natural tendency to assume that others "know what they mean," and that if there is uncertainty, questions will be raised. The writer, in contrast, is quite unlikely to resort to elliptical phrases and throwaway lines like "you know what I mean...." The writer, once having decided on the audience for a piece of text (or a speech), carefully designs the message so that it remains fairly constant regardless of the context in which the text is read. A friendly conversation, on the other hand, is likely to make sense only in the situation in which the conversation takes place. In a typical conversation, much is left unsaid, the assumption being that shared context and common experience will fill in the blanks. This strategy works most of the time, and communication seldom leads to serious emergencies. The situation can be quite different when the parties are the flight crew of a modern airliner, and their task is to ensure that hundreds of people make their way safely through the skies from one city to another. Under these circumstances, explicitness can become literally the difference between life and death.

The point of these remarks is the observation that the notion of explicitness and freedom from contextuality are intimately related. To the degree that the immediate situation provides the information needed for communication, the natural thing to do is to say only what needs to be said. Books cannot respond to questions, and so the writer must try to foresee any and all questions that might be raised by the audience, answering those as best as possible in advance. The text cannot include everything, and the writer always assumes prior knowledge on the part of the reader. The structures of assumed knowledge tend to be highly predictable, and it is the acquisition of these stylized schemata (Calfee 1981) for handling formal communication that is one of the most important outcomes of schooling.

Every child encounters a significant shift in style when leaving home and entering school -- if the teacher provides instruction in the formal use of language. Families by their very nature share many experiences in common -- the early development of language is an inherently natural phenomenon. The individuals who populate a classroom generally have much less in common, and the curriculum focuses on activities and events that tend to be relatively unfamiliar and abstract -- and properly so. Youngsters in the primary grades have had varying amounts of exposure to formal language prior to school entry. Some have learned to read and to value reading, whereas others have had less opportunity and encouragement in dealing with the printed word. Those parents who have had the advantages of a good education are most

NATURAL LANGUAGE (Utterance)	FORMAL LANGUAGE (Text)
Highly Implicit - Interactive	Highly Explicit
Context Bound	Context Free
Unique, Idiosyncratic, Personal	Repeatable, Memory-Supported
Intuitive	Logical - Rational
Sequential - Descriptive	Expository - "Content"

Figure 1. Contrasts between natural and formal modes of language and thought (Calfee, 1982).

27

likely to "talk like books." Real though these differences may be, virtually every child encounters in the school environment an emphasis on formalization that is alien compared to the comforts of home. School tasks have to be performed in the absence of strong contextual supports. Tasks often serve no immediate purpose, and meet no immediate obvious need. The students feel that school is a strange place -- and they are right. School is the place where we learn "unnatural acts" -- like reading (Gough & Hillinger, 1980).

Teaching for Transfer

A third concept underlying the design of the study is the importance of teaching for transfer, and of taking advantage of previous knowledge as a basis for transfer. At least one version of the bilingual hypothesis can be framed in terms of transfer of knowledge -- by teaching students to read in the language which they are most familiar, a foundation is laid for transfer of the basic principles of literacy to other languages, such as English. Throughout the design of the study, provision has been made to assess the extent of interlingual transfer of knowledge about language and reading, and about the degree of parallelism in the instructional opportunities provided in English and Spanish.

Transfer has significance beyond the bilingual hypothesis, of course. It appears that students are often expected to manage the transfer of knowledge on their own. Some students may be able to achieve this goal on their own, but it would seem more prudent instructional planning to ensure that students come to see relations and principles. Variation by the teacher in the instructional role (sometimes direct instruction, sometimes support), in the technique (sometimes analyzing, sometimes synthesizing), and in the kinds of materials (sometimes basal texts, sometimes library books, sometimes a class newspaper) may all be important indicators of instruction that promotes transferable skills and knowledge.

Brown, Campione and Day (1981) have pointed out the importance of metacognition in transfer -- students who understand what they know, who can explain to someone else what they have learned, are better able to apply this knowledge to novel situations. The free-form character of the reading assessment used in the study provides several opportunities for evaluating students' ability to articulate clearly their understanding of how to define words and recall text. In addition, specific metacognitive questions about each of the major components of reading are incorporated in the instrument. Student performance on these tasks should provide significant information on the extent to which students have acquired transferable knowledge.

Design Principles

The design of the study can be most conveniently separated into two segments -- the sample of districts, schools, teachers and students, which was determined by empirical considerations; and the design of the instrument package, which was determined primarily by theoretical considerations.

Sample Selection

The key concept in the design of the study is the identification of significant "natural variations" in existing programs. We were interested in looking at pre-existing variations in instructional programs between schools, the study of which could provide insights into program impact on children of differing learner cnaracteristics.

The overall design consists of three subcomponents: (1) the generalizability design; (2) the program design; and (3) the teacher/school design. The number of units selected at various levels of analysis (e.g., school district, school, teacher/classroom and student) was based on the anticipated distribution of available degrees of freedom among the three subcomponents of the design. The sample design is discussed in detail in a subsequent section, Sample Description.

Instrument Development

A number of instruments were developed or modified for use in the study. The instruments will be described in more detail in a later section; here the designed principles that guided the preparation of the instrument package is discussed.

A major consideration in the design was the creation of a package of instruments that reflected in a coherent fashion the concept of separable processes discussed earlier. In particular, the instruments for assessing reading achievement and reading instruction were planned so that the compo ts of decoding, vocabulary, and comprehension could be independently assessed, as could the students' and teachers' tendency to handle various facets of literacy and literacy instruction in a manner that was relatively more natural or formal.

A second major concern was the development of a valid instrument package. This requirement meant an emphasis on performance-based measures, o' en through individualized testing, and the creation of multiple measures for each of the major components. Reading achievement was therefore assessed by an instrument (the Interactive Reading Assessment System - IRAS, Calfee & Calfee, 1981), which was itself constructed according to multi-method design principles, but which was also backed up by a variety of other sources of information, including standardized tests and informal reading inventory data. Classrooms were observed on a regular schedule, but teachers were also asked to explain their intentions and the meaning of what was being observed, and were asked to discuss their lesson plans in a format that matched the design of the observation schedule.

The third principle entailed a careful analysis of the underlying curriculum structure, and attention to the representation in the achievement battery of developmental patterns within the curriculum. For instance, in the early grades, decoding is a major hurdle for yourgsters; requiring the kindergartener or first-grader to decode written language for assessment of word knowledge or comprehension means that the tester may not discover much about what the youngster knows in these areas. Accordingly, students were tested in parallel fashion on the ability to decode and to define words of varying familiarity, and on their skills at comprehending passages that they had to decode on their own, or that were decoded by the tester while they "read along." As another example, the first few grades of the elementary reading curriculum stress narrative text (stories); by third grade and increasingly thereafter, students must learn to deal with expository text (the style of writing encountered in areas like science and social studies). A third-grade student may decode at the fifth grade level, but still lack skills in comprehending the kinds of writing encountered in the uppe elementary grades. Accordingly, the assessment of comprehension included parallel passages representing both the narrative and expository genres from the second-grade level upwards.

The design principles laid out above were the basis for the fourth principle used in the development of the instrument package -- the creation of parallel instruments in both English and Spanish, and provision in the analysis of the instructional program for examination of English, Spanish, and bilingual components. In particular, the instruments used for assessment of Spanish reading achievement were not simple translations of the English instruments. Instead, a reading expert fluent in Spanish and knowledgeable about instruction in Spanish reading used the design principles and the general framework of the Interactive Reading Assessment System to create a parallel version of the instrument in Spanish.

The presence of a clearly articulated set of design principles has been vital in maintaining a high degree of coherence throughout the implementation of a research plan of extraordinary complexity:

- The research plan is both longitudinal and developmental. Some of the students have been followed for four years or more. Mobility and change in teacher assignments introduced complexities in data collection and analysis. Simply maintaining relations with a school site -- administrators, teachers, students, and parents -- over such an extended period of time has been a challenge.

- The project entailed dealing with two languages, each complex in its own right, and with approaches to instruction that vary across cultures as well as within cultures.

- Multiple cohorts were included in the plan. While this strategy provided a practical solution to some problems (e.g., the need to try out the instrument package, and to refine the techniques of data collection), it raised some problems in its own right (e.g., variations in details of the instrumentation, and extension of an already long time course for the study).

o.. Multiple sites were essential for generalizability of the findings, but they greatly complicated the task of data collection. The remoteness of some sites and the size of other sites are but two of the problems that had to be dealt with by the research staff.

o.. The study would have been much simpler if the design had relied on standardized measures of reading achievement. Most of the schools employed such tests for program evaluation, at least during some grades and for some students. On the other hand, not all students were administered these tests, nor was it altogether clear to the research staff that these instruments were entirely appropriate for assessing reading and formal language skills for the population investigated in this study.

o.. "Missing data" are always a problem in any large-scale research project, and are especially troublesome in a longitudinal study. Students may leave the neighborhood, either permanently or for a short while. Testers make mistakes in following procedures whenever these are relatively complex and require judgment, as was true in this study. Tape recorders malfunction. In some studies, the researcher may simply delete from the record any cases which include one or more missing data elements. That strategy was inappropriate to the present study for two reasons. First, the loss of a data element is often not a random event; certain situations and certain students are more likely to be associated with data loss, and it is frequently important to learn as much as possible about these students from the information that is available. Second, because of the extensive amount of data available for each student, one can properly view the study as comprising a large number of "case studies" -- when conducting a case study, one is ill-advised to drop the case just because a small amount of information is lost. Rather, the investigator usually takes advantage of existing information to make estimates of the missing data. When a reasonably coherent design is guiding the research, such estimation if often quite straight-forward -- this was the approach used throughout the study.

The Data Base

Data Structure

Because of the considerable extent and complexity of the study, it may be useful to provide a graphic representation of the overall data structure. Figure 2 shows the major elements and the specific sources for information within each element. The sources are listed in order of priority. Those at the top of each list are considered the most

PRECURSORS

Language Samples
Stanford Foundation Skills Test
Oral Language Proficiency Tests
Teacher Ratings

Cognitive Style Indices
Cartoon Conservation Scales

INSTRUCTION

Reading and Mathematics Observation System
Teacher Checklist
Attendance

Inventory of Bilingual Instruction
Survey of Teacher Language Skills
Teacher Cognitive Style

ACHIEVEMENT

Interactive Reading Assessment System
Standardized Tests
Informal Reading Inventory

Figure 2. Data structure for the SEDL Bilingual Reading Study.

valid informative data sources; those lower in the list are either less trustworthy, or are more difficult to quantify and/or interpret. The ancillary sources are of potential utility for cross-validation of the primary sources of information, though in some instances the validation may go in the other direction. Standardized tests, for instance, are widely used to draw inferences about the students' reading levels; the IRAS profiles have provided some useful insights into the reading components that are actually tapped by the standardized measures of reading.

Data Management

Seen from a distance, and presented within a broad framework such as the representation in Figure 2, the data have a neatness and precision that is quite appealing. For better or worse, the "real thing" is somewhat less orderly. Several procedures were established during the course of the study to help bridge the gap between the realities of the process of data collection in a field setting and the requirements of relatively "clean" data for the purposes of description and analysis. These procedures are summarized below.

First, it should be noted that considerable effort was given to monitoring the data collectors in the field to ensure that errors and misunderstandings were kept to a minimum. Nonetheless, the complexity and extent of the data base for the study guaranteed that mistakes would be made, that data would be missing, and that unusual events would occur. In addition, some of the instrumentation is quite innovative, and it was important to establish the reliability of the scales -- for that matter, it was also important to establish the trustworthiness of certain instruments that were presumably already validated.

The next stage in the data management process was the entry and cleaning of the data. Standard statistical packages were used to obtain descriptive summaries for identifying outliers and inconsistencies. For each item on each of the high-priority instruments, any variation (missing information, ambiguous data, observer mistakes, irregular patterns) was closely examined -- the original records were rechecked, observers were questioned, and where necessary, estimates were generated to handle missing or mistaken data at the item level.

As noted above, inter-item consistency was determined for each of the instruments used in the study, insofar as possible. The results of these analyses can be found in subsequent volumes of this report which deal with specific data sets. The inter-item reliability was used to ensure that the total-score measures for each subscale of each instrument were stable. Where necessary, items to be used in future testing were altered or changed to enhance the stability of the total score. These changes raised difficulties in maintaining the comparability of some of the scales over time -- the improvements in the instrument package were quite worthwhile, but how to equate scores from earlier versions to the results from the "new" and "improved" model? Such problems have been handled by a variety of techniques, which are described in subsequent volumes of this report in which data for each of the affected instruments are presented.

In some instances, an entire score was lost for a particular student or teacher. In these cases, correlative information has been used to estimate the missing information. A similar approach was employed to handle obvious outliers in the data. The estimates were based on "local" information, on other pieces of data that were substantively related to the flawed score -- estimates were not made in any instance to improve the efficiency of prediction equations for linking various data elements, but only to improve the internal coherence of separable data elements. For instance, if the listening comprehension section of the measure for a particular student was lost because of failure of a tape recorder, it seemed reasonable to estimate the score as at least as high a level as the student's reading comprehension score. If a student passed the vocabulary test at level D in second grade, and then performed only at level B at the end of third grade, and was reported by the third-grade tester as "having been distracted because the class was going on a field trip," it made sense to check other parts of the third-grade record, including other IRAS scores and the informal reading inventory, to determine whether the third-grade vocabulary score was a valid indicator of achievement.

The goal of the data management procedure has been the creation of an integrated data base over cohorts and over years, such that for each individual student a series of measures exists that is in all ways congruent with the series for every other student in the study, with the important stipulation that the number of missing data and outlier points be kept to an absolute minimum. The procedures necessary to achieve this goal have been perfected by the SEDL staff over the past several years, and the goal has been accomplished. The major tasks that were carried out during the last phases of the data collection period were (a) to incorporate the ancillary measures into the overall data base and (b) to enter the final data sets that were obtained during the last months of the final data collection year.

Data Reduction

The strategy for reduction of the study's complex data set to a manageable size was a straightforward application of the principle of "divide and conquer." The first step in the strategy has already been described -- the use of the inter-item consistency technique to obtain a set of reliable total scores for each subscale of the instruments used in the study.

The next step in the reduction process was guided by the theoretical concepts underlying the design of the study. The reduction process began within each of the major elements of the data structure shown in Figure 2 (see page 11). The Achievement segment can be used to illustrate the point. The output of the data-management procedure described above is a set of scores for IRAS, along with measures from the standardized tests and the informal reading inventories. Taking all facets of these instruments into account, the set of measures generated within the Achievement segment numbers some two dozen separate indices. One approach to handling this rather substantial set of summary measures might be to carry out a factor analysis. This approach assumes that we lack any theory about the structural characteristics of the data (which is not the case); moreover, the results of such an analysis are likely to be strongly influenced by the degree to which various facets of the reading process are represented in the set of indices.

A more reasonable strategy, from our point of view, was to bring together all those facets designed to measure similar components of reading achievement, first within a single instrument, and then to combine facets over instruments when justified and informative. Within this process, we decided to eliminate some of the data sources that simply were not informative. For instance, the performance of the students on the Letter Matching subtest of the pre-reading measure showed that virtually every child had adequate visual perception skills; despite the "Sunday supplement" stories about children who cannot recognize the difference between abstract symbols like printed letters, the data from the study replicate other findings that negate such claims. The various measures of cognitive style also deserved critical examination, in our opinion. Our examination of these measures for the data from the first two cohorts suggested that these indices may be of limited usefulness as precursors of success in reading and in response to instruction.

The goal of the data reduction strategy is reflected by the layout in Figure 2 (see page 11). We have worked to reduce the data structure shown in the figure from a set of three elements with 20 to 50 measures each to a set of three elements with 20 or fewer relatively independent indices, where the relations among indices from the three elements can be specified a priori in many cases.

Final Stages of Analysis

The primary analyses of the data from the study aimed toward four basic outcomes:

o.. Class-level descriptions of the approaches used to teach reading to children from bilingual backgrounds in Texas.

o.. Descriptive information using validated precursor profiles typically found in bilingual children on entry to schools throughout the state.

o.. Development and validation of a set of longitudinal achievement indices that could be used to assess growth in the various components of reading for English and Spanish.

o.. Development and validation of a set of procedures for measuring the linkage between reading achievement on the one hand, and precursor and instructional indices on the other hand, taking into account the possibility of interactions between precursor profiles and response to type of instruction.

Each of these outcomes is discussed in detail in a subsequent volume of this report: Measurement of Growth.

Sample Description

A number of interrelated factors were considered in selecting sites for this study, utilizing a purposive rather than a probability sampling procedure. The first consideration was to focus the study on the largest language minority group in the six-state region served by SEDL (Arkansas, Louisiana, Mississippi, Oklahoma, New Mexico, and Texas). Hispanics, whose backgrounds are tied to Mexico (Mexicans and Mexican Americans), constitute the largest and fastest-growing language minority group in the SEDL service region. Two of the states, Texas and New Mexico, rank among the ten states most active in bilingual education (third and eighth respectively) as indicated by the level of ESEA Title VII funds allocated to such programs. In the state of Texas, approximately one third of the children in public schools (872,000 students) are from Spanish-speaking backgrounds; many are limited in their English language skills. Similarly, New Mexico's public school population is heavily Hispanic. Approximately 61% of the children in grades one through three are from non-English-speaking backgrounds and are provided special language assistance programs. Of the 36,000 students in these programs, the large majority is Hispanic. In the other four states served by SEDL, clusters of Hispanic school children are identifiable, but in considerably smaller numbers than in Texas and New Mexico.

A second factor considered in site selection was the distribution of the target population within the SEDL service region. Given the large number of Hispanic students in the state of Texas alone (approaching one million), and the wide variation in the types of communities where these students live and attend school, the purpose of the study could be effectively and cost-efficiently accomplished by sampling sites within the geographical area encompassed by the state of Texas. Therefore, the primary sample population comprised children from Spanish-speaking backgrounds in the public schools of the state of Texas. A small sample of monolingual Spanish-speaking students, to serve as a comparison group, was also included from a region in Northern Mexico.

The sampling plan for the study included sampling at various units of analysis: region, school district, school, teacher/classroom, and student. The general approach employed was to start at the highest level of this chain with the selection of regions, and proceed to sampling at lower levels, using the best readily available data at each point to establish fixed categories from which samples were to be taken (Cronbach, 1976). Data compiled by the Texas Education Agency and in previous work carried out at SEDL suggested that two or three general types of bilingual education programs could be identified with two or three reading approaches nested within, or across, the bilingual programs.

Site Selection

The term "site" refers to the general community or school district within which a portion of the study was conducted. The initial selection of sites was based on the division of the State of Texas into geographical regions which took into consideration a combination of regional, political, and socioeconomic status, language, and degree-of-urbanicity variables (see Figure 3). Following discussions with members of the Texas Education Agency and educators throughout the state, four geographical regions (three in Texas and one in Northern Mexico) were identified:

Central Texas--a region which is both urban and rural and contains a number of bilingual programs;

Texas Border Area--rural, low socioeconomic status, substantial numbers of Spanish-dominant students;

East Texas--large urban area, largely monolingual English, middle socioeconomic status;

Northern Mexico--monolingual Spanish, rural and small and middle-sized cities, poor and middle class.

A cluster of bilingual programs also exist in rural and small city areas in the northwest Texas panhandle. This region was not included due to limitations of the study's fiscal resources.

The first three regions listed above constituted the primary regions from which the bilingual sample was drawn, while the last was a secondary region from which a comparison of monolingual Spanish-speaking children were sampled. Small groups of monolingual English-speaking children were also selected for comparison groups in the Central and East Texas sites.

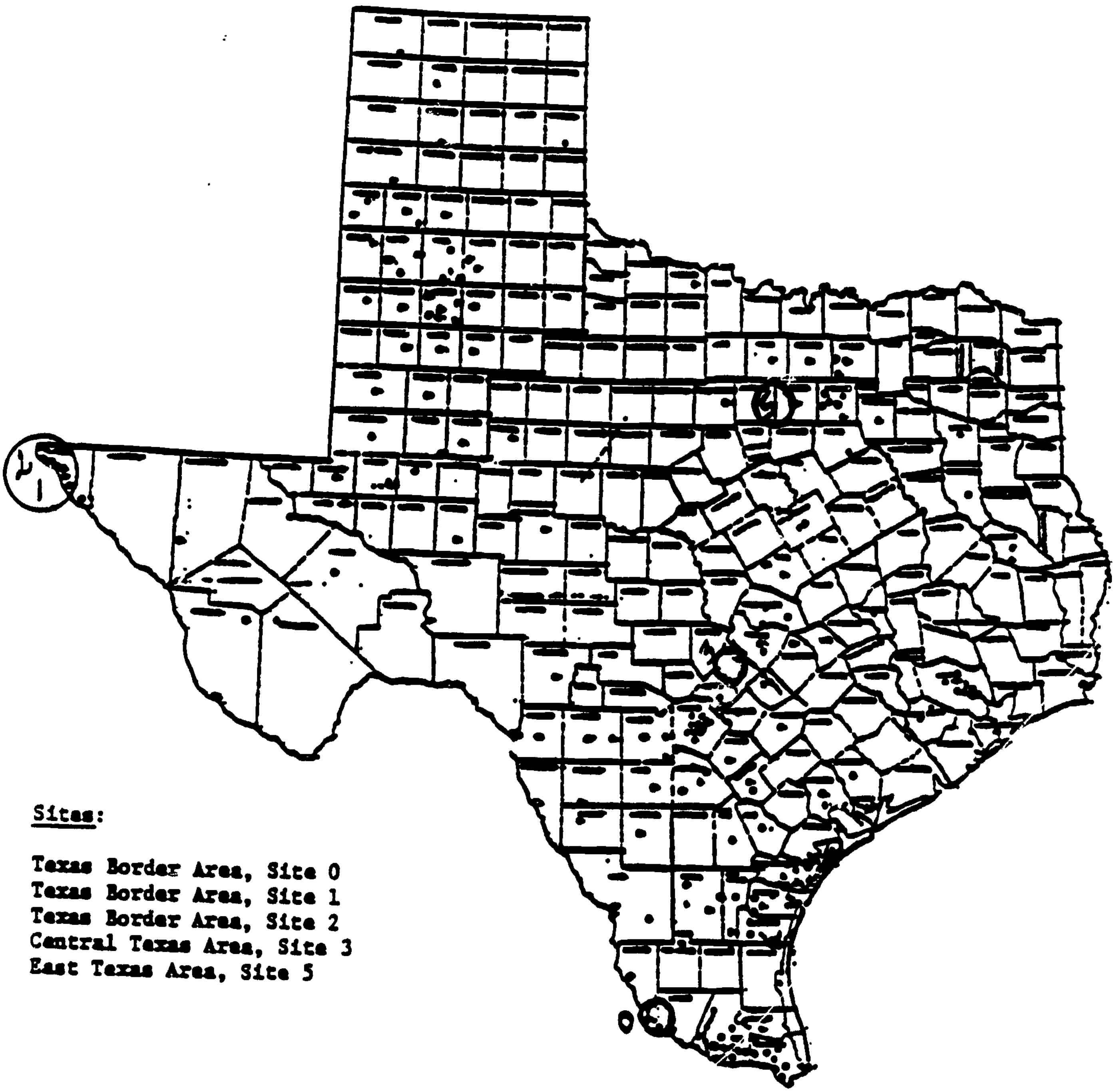
Selection of School Districts

Within each region, four to eight school districts were identified for potential inclusion in the study. Texas Education Agency data were used to gain information about the size of the districts (number of schools, teachers, and students), demographic profiles of the communities in the district (socioeconomic status, degree of urbanicity), and level of support per average daily attendance. Data were also available on established bilingual programs in each of the regions in Texas (Zamora, 1977). Following interviews with district personnel, six school districts were selected that were as broadly representative as possible on the variables of interest (see Figure 3). The location of the five Texas school districts selected for inclusion in the study are shown in Figure 4.

37

REGIONS	DISTRICTS	SCHOOLS	TEACHERS	STUDENTS
. Degree of urbanicity (urban, suburban, rural)	. Size (number of schools, teachers, and students)	. Size (number of teachers and students)	. Background-training	. Language proficiency
. Number of bilingual programs	. Socioeconomic status of communities in district	. Socioeconomic status of local community	. Nature of the reading program	. Cognitive development
. Concentration of Spanish-dominant students	. Degree of urbanicity	. Degree of urbanicity	. Nature of the bilingual program (primary region only)	. Reading achievement
. Range of socioeconomic status of communities in region	. Level of support per average daily attendance	. Level of support per average daily attendance	. Nature of students in teacher's class	. Cognitive style
. Geographical representation	. Bilingual programs (primary regions)	. Bilingual programs (primary regions)	. Willingness to cooperate in study	. Sex
. "Political" atmosphere	. Monolingual programs (secondary regions)	. Monolingual programs (secondary regions)	. Teacher's goals for reading	. Within classroom age
	. Nature of reading program(s)	. Nature of reading program(s)	. Time allotments for reading	. Mobility
	. Willingness to cooperate in study	. Nature of student body	. Time allotments for bilingual instruction	
	. Degree of variability of schools and teachers within district	. Willingness to cooperate in study	. Number of years of experience	
		. Character of the school organization (multiple grading, team teaching, open classroom configurations, or individually-guided education)	. Specialized training in reading	
		. Degree of within-school variation in use of reading methodologies	. Specialized training in bilingual education	
		. Configuration of teaching practices over grade levels K through grade four	. Attitudes toward multilingual children	
		. Teacher mobility	. Number of years at present school	
		. Degree of variability of teachers within school	. Qualifications and role functions of the teacher's aide	

Figure 3. Variables used in site selection for regions, districts, schools, teachers, and students.



<u>Sites:</u>

Texas Border Area, Site 0
Texas Border Area, Site 1
Texas Border Area, Site 2
Central Texas Area, Site 3
East Texas Area, Site 5

Figure 4. Schools with bilingual education programs - 1979-80 and location of the five sites in the SEDL Bilingual Reading Study (Northern Mexico, Site 4, lies south of Sites 1 and 2).

Selection of Schools

The data available suggested that schools are a major locus of program differences in the teaching of reading in bilingual programs in Texas (Zamora, 1977). Bilingual programs were installed in a number of schools throughout the state of Texas in an explicit form, while other schools did not have such a program. Also of interest was the variation between schools in the nature of the reading program per se. Consideration was given to these factors in the selection of schools within a district. Also considered was variation in the community socioeconomic status within the district, in the urban-suburban-rural status of the school, and in the character of the school organization. More specifically, this latter variable included such organizations as multiple grading, team teaching, open-classroom configurations, and individually-guided education programs (see Figure 3, page , for school selection variables).

Generally, schools were selected that had at least two teachers at each grade level. This was anticipated to provide a wider range of variation in teacher/program characteristics as the students were assigned to classrooms in subsequent years. However, in some instances a school with a single class at most grade levels was selected in order to reflect the realities of rural schools.

Fourteen schools were selected from which the student sample was initially drawn. As the students moved through the grades, six additional schools were involved. This resulted from bussing to achieve integration in one district. In other instances, the children proceeded beyond the grade level(s) served by their entry school and were, therefore, assigned to the school in their attendance area that normally served higher grade levels.

Selection of Teachers/Classrooms

Data available at the district and school level were used in the selection of teachers. Variables considered in teacher selection, as students entered the study initially, included number of years of experience, specialized training in reading and bilingual education, number of years at the present school, qualifications and role functions of the teacher aides (see Figure 3, page 17). As students moved on to the next grade, they were often disbursed throughout all appropriate classes that were available in their school at that grade level. Initially, the student sample was assigned to 26 homeroom classes. However, because of team-teaching and other organizational approaches, 37 teachers constituted the initial teacher/classroom sample. The distribution of this sample over the five years of data collection is shown in Table 1.

The goal of generalization underlies the selection of school/teacher configurations. By introducing a high degree of variability in program and teacher characteristics in the selection, it could be determined with some degree of confidence the range of conditions under which the findings of the study will hold. If, on the other hand, we

Table 1
Summary of Sample of Sites, Schools, Teachers, and Target Students by Grade Level of Students

Site	Grade	Schools	Teachers	Students
Border Area, Site 0	K	2	7	46
	1	2	12	62
	2	2	9	52
	3	2	11	54
	4	3	13	38
Border Area, Site 1	K	1	1	11
	1	1	5	22
	2	1	4	17
	3	1	4	15
	4	1	4	5
Border Area, Site 2	K	1	2	21
	1	1	4	40
	2	1	5	36
	3	2	4	31
	4	1	4	15
Central TX Area, Site 3	K	1	14	104
	1	1	16	93
	2	1	15	86
East TX Area, Site 5	K	6	11	143
	1	6	22	112
	2	5	21	84
	3	1	-	1
Northern Mexico, Site 4	1	2	5	54
	2	2	9	45
	3	2	6	35
Totals	K	13	35	325
	1	13	64	383
	2	2	63	320
	3	7	25	126
	4	5	21	58

42

had selected schools that were quite homogeneous in the makeup of their teachers and instructional programs, the findings could be generalized only to schools with similar characteristics.

Selection of Students

As will be noted in later sections, the students' language and reading skills were assessed with a variety of instruments, and their instructional programs and classroom instruction was observed and documented. For some purposes, all of the students in a class were tested with certain instruments; for other purposes, the instructional program for the entire class was observed. In addition, a target sample of 10 students was selected in each class for a more detailed "case study" examination. This target group of students was the subject of special observation and of individual assessment. They constitute the primary sample for this study.

The primary factors for the selection of target students within a classroom included sex, language status, and an index of cognitive style. Work that was underway at SEDL at the time at which the study was begun (De Avila & Duncan, 1979), as well as work carried out elsewhere (Lesiak, 1970; Stone, 1976), suggested that, for the purpose of this study, two traits ("field dependence/independence" and "reflectivity/impulsivity") could be used to summarize "cognitive style." A number of studies had indicated that cognitive style, as defined by either of these two traits, was a fairly good predictor of achievement in the lower grades, which is consistent with findings from earlier studies that indicated that cognitive style seems to have its greatest impact during the initial learning of particular skills.

For the purpose of target student selection, three measures were obtained for all students in each target classroom: language status, field dependence/independence, and reflectivity/impulsivity. The instruments used to obtain these measures are described below.

Language Status

The students' language status was determined on the basis of ratings made by the teachers for each of their students on the Student Operational Language Assessment Scale-SOLA (Duncan & De Avila, 1976) during the first month of school. Using this instrument, the teachers made an overall judgment about each child's language status on the basis of seven descriptions of language usage and skill; one was selected that best defined the student's observed ability to use Spanish and English: Monolingual English; Monolingual Spanish; Partial Bilingual - English Dominant; Partial Bilingual - Spanish Dominant; Bilingual (totally fluent in both English and Spanish); Limited English/Limited Spanish; Late Language Learner. This rating provided an impressionistic, global view of the child's ability in both languages.

Cognitive Style

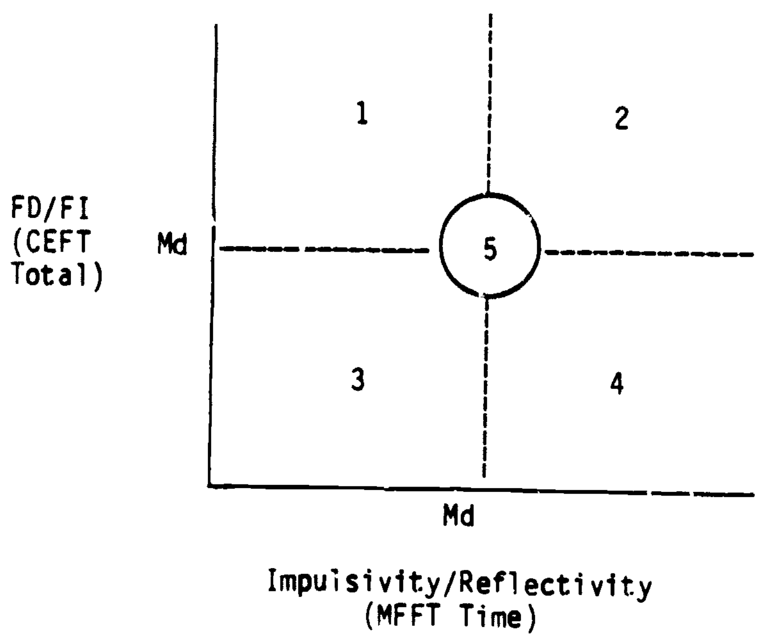
Two measures were used to assess "cognitive style." These were administered during the first month of school. The Children's Embedded

Figures Test-CEFT (Witkin, Oltman, Raskin, & Karp, 1971) is a measure of field dependence/field independence (FD/FI). The test, individually administered, requires the child to find a simple figure (a "triangle" or a "house" shape) embedded in a complex drawing (of a clown, for example). Credit is given for each time the student finds the simple figure without help from the test administrator. In order to reduce frustration, the child is helped to find the embedded figure if he encounters difficulty in a given item. The score is derived by adding the number of times the child finds the figure without assistance. Higher scores are associated with "field independence;" lower scores are associated with "field dependence."

The Matching Familiar Figures Test-MFFT (Kagan, 1965; Kagan, Rosman, Day, Albert, & Phillips, 1964), a measure of conceptual tempo (reflectivity/impulsivity), presents the child with a standard drawing of a figure (e.g., a ruler) and six variants. One is identical to the standard, with the other five differing from it in some detail. The child's task is to indicate which of the six alternatives is identical to the standard. The conceptual tempo classification is based on the time it takes the child to indicate the first response (latency) and the total number of incorrect responses made on each item (up to five errors per item). A child who is fast to respond (impulsive) often makes several errors; a child who is slow to respond (reflective) generally is more accurate.

Selection Procedure

Two-dimensional scatter plots of the students' performance on the cognitive style measures was developed for each target classroom, with CEFT defining one axis and MFFT (time) the other. Median splits were then used to divide the plots into four quadrants, with an additional subdivision being formed by drawing a circle of a specified radius at the intersection of the two medians, as shown below:



44

The five subdivisions defined five levels of the variable of cognitive style, a described below:

1 - field independent/impulsive
2 - field independent/reflective
3 - field dependent/impulsive
4 - field dependent/reflective
5 - moderate on both dimensions

Two target students were drawn from each subdivision, giving a total of 10 target students for each classroom. An effort was made to balance, to the extent that it was possible, the students selected in each subdivision (across classrooms) on the variables of sex and language status. Our purpose in the selection was to maximize the systematic variation between students in order to achieve a basis for generalization of the results. The total number of students involved in the study and their distribution among the sites by grade level is shown in Table 1, page 20. The student sample, by data collection yea~, is discussed in further detail in a subsequent section of this document (see page 30).

Characteristics of the Sites

The following description focuses on the characteristics of the target district at the time of selection. First, demographic data, relative to the Texas sites, are presented which characterize the size, socioeconomic status, degree of urbanicity, local funds available per student, bilingual program(s), concentration of Hispanic students, and variability of schools in each district. Second, certain variations in reading programs across the five sites are summarized, focusing on type of instructional materials used, organization for instruction, criteria for transition from Spanish to English, and instructional emphasis. Third, the extent to which the selected districts are representative of other districts in Texas is examined in terms of size, concentration of Hispanic students, and location of bilingual programs. Lastly, a brief description of the Northern Mexican site is presented. A detailed description of each of the sites is included as Appendix A.

Demographics

As shown in Table 2, the five districts selected from the state of Texas to participate in the study varied considerably on each of the demographic indices. The urban district (East Texas Area--Site 5) is large, with a well-developed bilingual program, yet it differs substantially from the other districts in (a) locale, (b) percentage of Hispanic students, (c) per pupil expenditure, and (d) variability of schools within the district. The two middle-sized districts (Central Texas Area--Site 3 and Texas Border Area--Site 0) are very similar in terms of SES and degree of urbanicity, yet differ considerably in (a) locale, (b) percentage of Hispanic students, (c) per pupil expenditure, and (d) variability of schools within the district. The two smaller districts are, in fact, quite dissimilar in size (Texas Border Area--

Table 2
Demographics for Texas Districts in the SEDL Bilingual Reading Study - 1982

Variables / Site	SIZE: Schools	SIZE: Teachers	SIZE: Students	SES of Communities in District	Degree of Urbanicity and Locale	Per Pupil Expenditure	Bilingual Programs	Concentration of Hispanic Students: District-wide	Concentration of Hispanic Students: Elementary Schools	Degree of Urbanicity Within District
Texas Border, Site 0	8	281	6,460	88% low SES	Small town & rural areas, south Texas	$2,428	K-5 (ESL 6-10)	98%	97%	Some schools located in small town; others in rural, isolated areas
Texas Border, Site 1	2	40	798	low income SES	Very small & rural areas; far west Texas	$1,467	K-12 (ESL (K-12)	97%	98%	1 elementary school in district ([illegible] -K-6)
Texas Border, Site 2	4	103	1,625	low to lower-middle SES	Small town and rural areas; far west Texas	$1,964	K-6 (ESL K-12)	90%	90%	1 primary K-3; 1 elementary 4-6
Central Texas, Site 3	6	270	4,615	low to lower-middle SES	Medium-sized town; central Texas	$1,842	K-3 (ESL 4-12)	66%	66%	1 campus for each of 2 grade levels; K-1 is open-classrooms; 2-3 & 4-5 are self-contained
East Texas, Site 5	99	3,069	65,125	low income-44%; middle or upper income-56%	Large urban area; northeast Texas	$2,724	K-6 (ESL 6-12)	19.7%	varies; concentrated in 18 schools	26 schools have some type of bilingual programs

Site 2 has three time the enrollment of Texas Border Area--Site 1) and in amount of per pupil expenditure. They are alike, however, in terms of (a) SES, (b) locale, (c) concentration of Hispanic students, and (d) variability of schools within the district. Taken as a whole, these sites represent a broad range of communities in Texas where bilingual programs have been implemented, from the large urban district in northeast Texas to the small rural districts along the Mexican border. Thus, the educational processes and outcomes observed in this study should be representative of many, if not most, of the bilingual programs in the State.

Variations in Reading Programs

As shown in Table 3, the five Texas districts varied substantially at the time of selection in terms of their reading programs for bilingual students. The four variables examined were (a) materials, (b) organization for instruction, (c) criteria for transition from Spanish to English reading, and (d) instructional emphasis in reading. The information depicted in Table 3 came from three sources: documents provided by each of the districts describing their bilingual programs; interviews with administrators and teachers; and formal and informal classroom observations.

Materials. In two of the five districts, basal reading series provided the foundation for the English reading program. These series are structured so that the sequence of instruction is built into the readers and workbooks, which gradually increase in difficulty as the child progresses. Three districts utilized management systems. These are characterized by instructional objectives, mastery tests, and a diversity of materials (both commercial and teacher-made).

For the Spanish reading program, four of the districts relied on basal reading series for instruction. One district was employing an "individualized" approach to instruction that drew upon the "Guszak Diagnostic/Prescriptive Reading Program" (Guszak, 1972). In this program, management was carried out through a system of student contracts.

The materials used for reading instruction across the five sites represented the two curriculum approaches most prevalent in bilingual programs at the time of sample selection: basal series and management systems. As will be described later in this section, the instructional emphasis varied considerably within these two basic curriculum approaches.

Organization for Instruction. The organization for reading instruction in bilingual programs in Texas follows one of four basic patterns: (1) self-contained classrooms, (2) team-teaching, (3) open classroom structure, and (4) bilingual resource teacher(s). Each of these patterns was represented in the sampled classrooms in the study.

Self-contained classrooms, in which one teacher provides all (or most) of the instruction for a given group of children, is characteristic of most of the school districts in the State. This organizational

Table 3
Variation in Bilingual Reading Programs on District Entry into the Study

	Texas Border Site 0		Texas Border Site 1		Texas Border Site 2		Central Texas Site 3		East Texas Site 5		Northern Mexico
Materials:	English	Spanish	English	Spanish	English	Spanish	English	Spanish	English	Spanish	Spanish
. Basal Series	x	x	x	x	x	x	x	x	x	x	Federal textbooks
. Management System	x				x	x	x				
Organization:											
. Self-contained	x	x			x	x			K-1	K-1	x
. Team-teaching			x	x			x	x	Gr.2	Gr.2	
. Open classroom							x	x			
. Resource bilingual teachers										In some schools/classes	
Transition Based On:											N/A
. Spanish reading skill				x		x		x		x	
. Oral English proficiency			x		x		x		x		
. No transition; concurrent Spanish/English reading instruction	x	x									
Instructional Focus:*											
. Decoding	x	x					x				x
. Skills development			x	x				x	x	x	
. Meaning-based					x	x					

*Focus in early phases of reading; see text for explanation of terminology.

pattern existed in two of the border area sites, and in most of the schools in the urban district.

Team-teaching is characteristic of the instructional programs in the Texas Border Area--Site 1, Central Texas Area--Site 3, and in some of the schools in the urban district (East Texas Area--Site 5). In this organizational structure, one teacher typically provides Spanish reading and/or content area instruction for Limited English Proficient students, while the other member(s) of the team provides reading and content area instruction in English for those children for whom such instruction is appropriate.

The open classroom structure, found in the K-1 school at the Central Texas Area--Site 3, involves large, open spaces that house "units" of approximately 100 children in which some 8 teachers and 5 aides work together to provide individual and small-group instruction to children of the various language classifications. Typically, each team includes one or more bilingual teachers who carry out the Spanish component of the instructional program.

The use of a bilingual resource teacher to deliver instruction to Limited English Proficient students was found in only a few schools in the urban site. In this organizational pattern, one teacher provides Spanish reading and content area instruction, usually in small groups, for all Limited English Proficient students at a particular grade level or in a given school.

Criteria for Transition from Spanish to English Reading Instruction

In examining this variable, it became necessary to look at both policy and practice to get a clear picture of what was happening in each of the districts. Policy for all of the target districts provided for transition from Spanish to English instruction, with English-only instruction being the eventual goal. Practice varied, however, as to the time of transfer from Spanish to English, both within and between districts. Criteria for transition at all of the sites included both Spanish reading ability and English oral proficiency. According to the school personnel interviewed at the time of selection, some individual schools and/or teachers begin English reading instruction earlier than others. At one of the sites, English reading instruction was observed to begin for all children by the second semester of first grade. As documented during the course of the study, actual practice at the other sites varied considerably both within site and among the sites.

Instructional Emphasis

The instructional emphasis appeared to be determined, to a large extent, by the theoretical orientation of the teacher and by the nature and characteristics of the teaching materials available within the school district. School policy, as to the uniformity of the basic materials to be used and the approach to be followed. also affected the instructional emphasis. Similarly, the amount and kind (focus) of "current" inservice activities within the district influenced the instructional emphasis.

While varying from teacher to teacher, the instructional emphasis in two of the five districts selected for the study (Texas Border Area--Site 1 and East Texas Area--Site 5) could best be characterized as "skills development," in which the components of decoding, vocabulary development, and comprehension of whole text are given relatively equal attention during the early phases of reading instruction. These two districts had adopted specific basal reading series in both English and Spanish and had encouraged district-wide use of these texts.

In the English reading program in the Texas Border Area--Site 0, the "Wisconsin Design" management system was used which focused heavily on letter-sound correspondence and work attack skills in the early stages, with increasing attention given to comprehension skills as the child gained some facility in reading. The reading series adopted by the district also placed a strong emphasis on word recognition in the early stages. Similarly, the Spanish reading component utilized a basal reading series in which formal instruction in letter-sound correspondence and word recognition skills was emphasized. Comprehension skills received lesser attention in the early stages of reading instruction.

The English reading program in the Central Texas Area--Site 3 was similar to that of the Texas Border Area--Site 0 described above. However, the Spanish reading program differed in that this site utilized a basal reading series that emphasized comprehension skills, in the early stages of reading, to a greater extent than did the basal series adopted at the Texas Border Area--Site 0.

At the time of site selection the reading program at the Texas Border Area--Site 2 employed a management system locally referred to as the "Guszak Diagnostic/Prescriptive Reading Program." As noted earlier, this system is characterized by individualized instruction managed through the use of student contracts. Its theoretical orientation is strongly "meaning-based," with little or no formal instruction in letter-sound correspondence until after the child has gained some reading fluency. Two years into the study at this site, the district shifted to a basal reading program and abandoned their individualized approach in both English and Spanish for a more traditional small-group instructional procedure.

Summary

Whether considering the variable of curriculum materials used, organization for instruction, criteria and practices for transition from Spanish to English reading, or instructional emphasis, the reading programs of the five Texas districts selected for the study reflect a high degree of diversity. Thus, the naturally-occurring variations necessary to the design of the study are found in the sites included in the study.

Representativeness of Districts Selected

The five districts selected represent a cross-section of school districts typically found in the State of Texas. The following discussion examines the degree to which these districts are comparable to other districts in the State in terms of size, concentration of Hispanic students, and distribution of bilingual programs.

Size

The Texas Education Agency (1976) divides the districts into 12 categories according to size. The first category comprises six districts, each with more than 50,000 students. These six metropolitan districts serve one fourth (550,000 students) of the two and a half million students in the State. One of these, East Texas Area--Site was included in the study. Approximately one fifth of the students in Texas are in the 189 middle-sized districts (1,500 to 4,999 students). Three of the five districts sampled for this study fall into this category (Central Texas Area--Site 3, Texas Border Area--Site 0, and Texas Border Area--Site 2). While the small school districts in Texas (less than 500 students) serve only about five percent of the state's school children, they represent approximately one half (541 out of 1120) of the districts state-wide. One site, Texas Area--Site 1, was selected to represent this category.

The sites selected for the study represent the full range of large, medium and small districts in Texas. They reflect both the size of district where most students are enrolled (metropolitan and middle-sized) as well as the size of district which characterizes most of the school districts in the state (small).

Concentration of Hispanic Students

In order to examine the representativeness of the sampled districts in terms of concentration of Hispanic students, the districts in the state were divided into three categories on the basis of data collected in the fall of 1979 by the Texas Education Agency (1980). They were classified by percentage of Hispanic students as follows: 80% to 100%; 60% to 80%; less than 60%. Seventy-two districts were found to have high concentrations of Hispanic students state-wide with three of the five districts selected for this study falling into this category (Texas Border Area--Sites 0, 1, and 2). Sixty-one districts fell into the 60% to 80% category, with one of the sampled sites (East Texas Area--Site 5) in this category. The remainder of the 1,099 districts responding to the Texas Education Agency survey had less than 60% enrollment of Hispanics; the Central Texas Area--Site 3, with 17% Hispanic enrollment, represents districts in this category. Thus, the five sampled sites represent the full range of districts where bilingual reading and program implementation are crucial concerns--those with high, medium-high, and lower percentages of Hispanic students.

Distribution of Bilingual Programs

As shown in Figure 4, page 18, the five sites are located either in regions where other bilingual programs are concentrated or are in a locale which is very similar to other districts with bilingual programs. Three of the sites are along the Texas-Mexico border, with one in south central Texas and two in southwest Texas. The site in central Texas represents stable Hispanic populations that have somewhat less contact with the Mexican culture and the Spanish language. The urban site in East Texas represents the wide dispersion of districts with bilingual programs in the northern and northwestern regions of Texas. Thus, the sampled districts represent (1) the border region; (2) the region 150 to 200 miles from the border; and (3) the region more than 200 miles from the border. These regions, in turn, reflect varying degrees of contact with the Spanish language and culture as well as differing degrees of concentrations of Hispanic students in the schools.

Northern Mexico, Site 4

Two schools, located in an isolated, middle-sized Mexican city some 200 miles south of the Texas-Mexico border, comprise the sample for this region. One of the schools is a state-supported school; the other is a federally-funded school. Both schools are located within the city boundary; both serve primarily monolingual Spanish-speaking students from low to lower-middle income families. At the time of selection, the children were attending school for four and one-half hours per day. The classes were large (approximately 50 children per class). The classes were self-contained, and all instruction was provided by a teacher who had completed normal school training. All of the classes were using the federally-adopted textbooks in which all subject matter is integr ed into one set of books (i.e., math, science, social studies, and language arts are interwoven into the same textbooks). Reading instruction per se focused heavily on letter-sound correspondence in the early stages; handwriting and composition were an integral part of the reading instruction, even at the early grades. Although there was variation in some of the classes in one of the schools, most of the instruction was directed to the full group and was characterized by much direct instruction on the part of the teacher and choral response on the part of the students.

Cohort Plan for Longitudinal Investigation

It was most desirable, in order to achieve the purpose of this study, to track the target students from entry into kindergarten through the end of fourth grade. The growth and development that are the focus of this study normally takes place over this time period, and a cross-sectional design would be altogether inappropriate. It should be emphasized that, for practicality, what was planned and carried out was the selection of cohorts of relatively modest sample size who were tracked for varying periods of time in successive waves.

Cohort 1

The first cohort, consisting of 12 classrooms in the Texas Border Area--Site 0, was selected for testing and observation during the 1978-1979 school year. Four of the classrooms were selected as target classrooms to remain in the study, and 10 target students were selected in each of these classrooms. The full range of instruments and data collection procedures which had been prepared for the study were thoroughly tested under specified conditions in all 12 classrooms, and modifications were made as needed. Thus, this cohort, which served as the pilot cohort, consisted of 40 students (20 kindergarten and 20 first grade students) from four classrooms in two schools (one kindergarten and one first grade from each school) in one school district in one region of the state.

Cohort 2

During the first data collection year (1978-1979), a second cohort of eight classrooms was identified for inclusion in the study in the school year of 1979-1980. Two of the classrooms (one kindergarten and one first grade) were located in one school in the Texas Border Area--Site 1 Four classrooms (two kindergarten and two first grades) were selected from one school in the Texas Border Area--Site 2, and two additional classrooms of kindergarten students were added from the original schools which contributed Cohort 1 in the Texas Border Area--Site 0. Thus, two successive waves of class/teacher/student cohorts were tested and observed, the second replicating the basic design of the first one.

Cohorts 3 and 4

In the 1980-1981 school year, Cohorts 3 and 4 entered the study. Cohort 3 consisted of students from the East Texas Area--Site 5 and then Central Texas Area--Site 3. Eleven classrooms were selected from one school district in Site 5: eight bilingual classrooms of kindergarten students in seven schools and three English-medium classrooms of kindergarten students from one school which housed one of the target bilingual classrooms. Thus, from those classrooms, a subsample of 80 bilingual children and 30 monolingual English-speaking children entered the study.

In the Central Texas Area--Site 3, 80 bilingual and 10 monolingual English-speaking kindergarten children were sampled from three team-teaching units which contained some 300 children within one school. Fourteen homeroom and language arts teachers were associated with this group of students.

Cohort 4 consisted of students from the Northern Mexico Area--Site 4. Four classrooms of first grade students in two public schools (two classrooms at each school) were selected with a subsample of 15 monolingual Spanish-speaking target students per class. Student attrition was expected to be higher in the Mexican schools than in the Texas schools, thus the selection of a relatively larger subsample at the point of entry was calculated to ensure an adequate sample of students over the course of the study.

Summary

Four student cohorts of differing size entered the study during the course of a three year period. Each cohort of students was tracked from their entry into the study through the last data collection year (1982-1983). The selection procedure yielded a subsample of 380 students distributed as follows:

Cohort 1 - 40 students (20 kindergarten; 20 first grade)

Cohort 2: 80 students (50 kindergarten; 30 first grade)

Cohort 3: 200 kindergarten students

Cohort 4: 60 first grade students.

Since the students entered in successive waves, with most of the students entering during their kindergarten year but with some entering at first grade, certain of the students were tracked for five years (K-5); others for four years (K-3 or 1-4); and yet others, which were the majority, were tracked for three years (K-2 or 1-3, the latter being the case of the Northern Mexico sample).

Once selected to participate in the study, a student was followed through whatever classrooms he was subsequently assigned. Such assignments were based entirely upon criteria established by each school district, and without any input from SEDL personnel. During the data collection phase of the study, a student remained in the study until he either (1) completed fourth grade, or (2) moved outside the schools involved in the study. On a few occasions, a student was dropped from the study at a parent's request. If a student left the study prior to first grade entry, a replacement student from the same initially-selected classroom was added to the study who possessed similar characteristics to the replaced student based on the initial selection criteria (i.e., same cognitive style quadrant, sex, and degree of bilingualism). For such replacement students, every effort was made to acquire a complete data set by collecting data from the school district files (e.g., the district-administered standardized language test scores), and by administering any SEDL assessments missed by the replacement target during the semester in which the student was replaced. If a student left the study after first grade entry, a replacement was not sought since it was felt that too much missing data would result for such replacement students. As noted earlier, the attrition rate at Site 4 (Northern Mexico) was expected to be quite high, and oversampling was employed to meet this attrition problem rather than following the replacement procedure.

In all, 438 students were identified as targets during the five-year data collection phase. The attrition rates varied over sites, being highest at Sites 1 and 5 (about 37% over the entire data collection period), lowest at Site 3 (18%), and averaging 28%. Similarly, the replacement percentages were highest at Sites 1 and 5 (24% and 30%, respectively), with the remaining sites at about 11%, for an overall average of 17%.

As shown above, the initial sampling plan called for the selection of 380 students. In Site 4, 60 students were to be sampled (15 from each of 4 classrooms), but due to the joint distributions of the cognitive style measures obtained from each of these classrooms, only 54 students could be selected according to the sampling criteria. At Site 0, a brother of a selected twin was added to the study for a potential case study analysis. Also, at 2 additional sites (1 and 3), a total of 3 initially-selected targets left the study, were replaced, but then returned to the study and were once again tracked along with their replacements. Given these sampling adjustments, the planned sample of 380 students resulted in an actual sample of 378 students.

Of the initially selected 378 targets, 105 students eventually withdrew from the study; 60 of these students were replaced and 47 of these replacements remained in the study until normal exit. Since the primary objective of the study was to track reading acquisition and our most extensive assessments of reading skills were begun at the end of first grade, only students who remained in the study through at least two additional years of instruction beyond kindergarten were included in the longitudinal analyses, thus providing us with at least two data points to chart growth in reading. A total of 333 students met this criterion.

All longitudinal analyses were based on instructional years rather than grade levels in order to track the number of years of actual instruction regardless of whether a student was retained or double-promoted. Such irregularities only occurred after first grade in our sample, so that instructional year 0 is always kindergarten, and instructional year 1 is always first grade; later instructional years (2-4) may be either a grade level above (for students double-promoted sometime after first grade) or below (for students retained sometime after first grade) the nominal instructional year value. No student in our sample was retained or double-promoted more than once.

A breakdown by cohort and language group of the longitudinal sample of 333 students is presented in Table 4. As can be seen from the table, 32 students (10% of the sample) had irregular grade sequences. Considering the 254 bilingual students which are of primary concern in this report, the table shows that 153 students were tracked through two instructional years beyond kindergarten (all but one having kindergarten data as well); 40 students were tracked through 3 instructional years (with all but two having data at kindergarten); and 61 students were tracked through four instructional years (with 36% having kindergarten data).

Instrumentation and Data Collection

Schedule of Testing and Observation

The purpose in tracking the students from their early school experience through the mid-elementary years was to trace the full development of oral language and reading skills. As students move from kindergarten through fourth grade they are exposed to a variety of instructional programs. Our goal was to track the reading progress of

Table 4
Breakdown of Instructional Year Sequence for Longitudinal Sample by Cohort and Language Group

Cohort	Language	01234	1234	0123	123	012	12	Totals
		Instructional Year Sequence						
I	Bilingual	16 (1)	19	1 (1)	1	1 (1)		38
II	Bilingual		26 (7)	37 (5)	1	5	1	70
III	Bilingual					146 (12)		146
Total		16 (1)	45 (7)	38 (6)	2	152 (13)	1	254 (27)
III	Monolingual English					36 (2)		36
IV	Monolingual Spanish				38 (3)		5	43
Total								333

Note: Numbers in parenthesis represent the number of students with irregular grade sequences.

students of differing learner characteristics and to observe their responses to variation in the instructional program. As will be explained in more detail in the section on instrumentation that follows, an ongoing program of observation and testing was conducted, such that variation in the instructional program over a period of a few weeks was highlighted in the data, and where the student was tested on a regular basis so that changes in student performance was also measured within these time frames.

Data were collected each year in accordance with a data collection schedule which was prepared each summer and distributed to data collectors prior to the beginning of the school year. In some cases, slight adjustments were made in the timetable to accommodate a particular school's planned activities or its emergencies (e.g., vacation schedules, school-wide testing, flu epidemics). However, in general data collection followed the sample schedules presented in Tables 5 and 6.

Instrumentation

This section presents a brief overview of the data sources for the study. One set of measures were used to assess student characteristics and academic performance; the other provided information on teacher characteristics and classroom instruction. While only a brief discussion of the instruments is presented here, a detailed description of each of the data sources is provided in the corresponding volume of this report where the results of the data analyses are presented.

Student Characteristics and Student Performance

Language assessment. Three types of language measures were used to assess the child's oral proficiency in both Spanish and English: (1) an oral language proficiency test, (2) teacher ratings, and (3) an ethnographic verification of the child's oral language abilities (audiotaped language samples).

Oral language proficiency test. School districts in Texas were required under the Texas State Plan for Bilingual Education (Texas Education Agency, 1978) to identify all children who are exposed to a language other than English outside the school environment, and to administer to each of them, on their initial entry into the school district program, the English version of one of the commercially-available oral proficiency tests approved by the state. Frequently, school districts also administer the Spanish version of these tests to Spanish-speaking children on their entry into bilingual programs. These language data were collected from the schools; where Spanish-language data were not available for students on their initial entry, the project staff administered to each of the target bilingual children the Spanish version of the oral language proficiency test selected by the district.

It was also necessary to administer both the English and Spanish versions of the district-selected test to each of the target bilingual children yearly in grades beyond kindergarten, since the State Plan did

Table 5
SEDL Bilingual Reading Study - Texas Sites
Data Collection Schedule: Grades 2-4

Month	Classroom	Time Mins.	Teacher	Time Mins.	Grades 2-4	Time Mins.
September			Interview I	30	Children's Embedded Figures Test (CEFT)	30
			Language Estimate #1	10		
					Matching Familiar Figures Test (MFFT)	20
					Oral Language Proficiency Test (Spanish & English)	40
October	Observation RAMOS/SEDL	0	Checklist (#1)	10	Informal Reading Inventory (IRI)	20
			Inventory of Bilingual Instruction			
November	Observation RAMOS/SEDL	0			Language Sample	0
December	Observation RAMOS/SEDL	0	Language Estimate #2	10	Language Sample	0
					Informal Reading Inventory	20
January	Observation RAMOS/SEDL	0	Checklist (#2)	10	Cartoon Conservation Scales (CCS)	30
			Survey of Teacher Language Skills	10	Language Sample	0
February	Observation RAMOS/SEDL	0			Language Sample	0
					Informal Reading Inventory	20
March	Observation RAMOS/SEDL	0	Checklist (#3)	10	Language Sample	0
			Embedded Figures Test	20	Interactive Reading Assessment System (IRAS) (Spanish & English)	60
			Matching Familiar Figures Test	10		
April/ May	Observation RAMOS/SEDL	0	Interview II	30	Language Sample	0
			Language Estimate #3	10	Standardized Test Scores	0
	Attendance Records	0				

Table 6
SEDL Bilingual Reading Study - Northern Mexico Site
Data Collection Schedule: Grade 3

Month	Classroom	Time Mins.	Teacher	Time Mins.		Time Mins.
September			Interview I Language Estimate #1	30 10	Children's Embedded Figures Test (CEFT) Matching Familiar Figures Test (MFFT) Oral Language Proficiency Test (Spanish & English)	30 20 40
October	Observation RAMOS/SEDL	0	Checklist (#1) Inventory of Bilingual Instruction	10	Informal Reading Inventory (IRI)	20
November	Observation RAMOS/SEDL	0			Language Sample	0
December	Observation RAMOS/SEDL	0	Language Estimate #2	10	Language Sample Informal Reading Inventory	0 20
January	Observation RAMOS/SEDL	0	Checklist (#2) Survey of Teacher Language Skills	10 10	Cartoon Conservation Scales (CCS) Language Sample	30 0
February	Observation RAMOS/SEDL	0			Language Sample Informal Reading Inventory	0 20
March	Observation RAMOS/SEDL	0	Checklist (#3) Embedded Figures Test Matching Familiar Figures Test	10 20 10	Language Sample Interactive Reading Assessment System (IRAS) (Spanish & English)	0 60
April/ May	Observation RAMOS/SEDL Attendance Records	0 0	Interview II Language Estimate #3	30 10	Language Sample Standardized Test Scores	0 0

not require formal retesting of the children for oral language proficiency subsequent to initial entry into the school district's program. The monolingual English-speaking children in the study were administered yearly the English version of the district-adopted test. The monolingual Spanish-speaking children were administered a Spanish oral language proficiency test selected by their schools.

Teacher ratings. On three occasions during the school year, the teachers provided their evaluation of the children's language usage and ability. As noted in a previous section on the selection of students, teachers in target classrooms were asked, at the time of selection, to rate all of the children in their classes on the Student Operational Language Assessment Scale (Duncan & De Avila, 1976). Target children in each of the classrooms in subsequent years were also rated by their teachers on this instrument. These ratings, along with other data, were used to construct an index of each child's language ability and growth over time.

In the month of December, after the teachers had become more familiar with the language ability and usage patterns of their students, all target children were rated by their teachers on the Oral Language Proficiency Rating Scale (Mace-Matluck, Tunmer, & Domínguez, 1979). The teachers were asked to rate the target children once again on this instrument during the month of April or May, concurrent with the administration of the reading achievement tests. These ratings provided the teacher's evaluation of specific aspects of the language ability of each child in both Spanish and English, as well as an overall, global performance rating.

Language Samples. Additionally, for the purpose of monitoring the child's language growth as well as for verifying the child's language status, audiotaped speech samples were obtained monthly from the target children (Mace-Matluck, Tunmer, & Domínguez, 1978). These were taken on a rotating schedule in three communication settings: in the classroom, in the home, and on the playground or in other non-instructional settings within the school.

Reading assessment. Assessment of the children's progress in reading was conducted on a planned schedule and involved four types of instruments: (1) a reading readiness measure, (2) an informal reading inventory, (3) a reading achievement measure, and (4) standardized achievement tests.

Reading readiness. The Stanford Foundation Skills Test/Prueba Stanford de Destrezas Fundamentales (Calfee & Associates, 1977; Calfee & Peña, 1978) was selected as the reading readiness measure for the study. This test battery is designed to assess reading-related skills, in Spanish and in English, of children in kindergarten and first grade who have not yet learned to read in any language. Included in the assessment are tasks of alphabet recognition, letter matching, phonetic segmentation, vocabulary distinction, common labels, and conceptual understanding.

64

All bilingual children that entered the study at kindergarten were administered the reading readiness instrument in both Spanish and English in the early fall, as were a subsample of target children at grade 1. Monolingual children were administered the version appropriate for them.

Informal reading inventory (IRI). Informal reading inventories were developed by the research staff in both English and Spanish (Mace-Matluck & Domínguez, 1978a, 1978b; Mace-Matluck, Domínguez & Padilla-Hajjar, 1978). All target children were administered an IRI at their entry into a formal reading program (i.e., instruction in connected reading that follows any "pre-reading" instruction that is typically referred to as reading readiness), and their reading progress was monitored, through the use of the IRI, on a monthly basis throughout the school year. Bilingual children were administered the IRI's in both English and Spanish, regardless of the language of the reading instruction provided in the classroom.

Reading achievement. The Interactive Reading Assessment System-IRAS was employed as one of the measures for assessing the student's reading ability. The IRAS, an individually administered diagnostic assessment system, was designed for research application initially and has undergone two revisions (Calfee & Calfee, 1979; 1981). Modeled after the informal reading inventory, the IRAS provides independent measures of several component skills essential for fluent reading. The materials in the instrument were selected to cover a wide range of skills and knowledge in the areas of oral language and reading from the level usually expected of a mid-year first grader to that of a junior high school student. The Spanish version of the IRAS was developed in 1979 using the same ormat and procedures used in the development of the English edition Calfee, Calfee & Peña, 1979).

The areas of skills and knowledge assessed in the system include: reading of isolated words, definition of common words within and beyond the student's reading vocabulary, and selected word analysis skills based on the pronunciation of synthetic words. Comprehension of connected text is also assessed in contexts of both narrative and expository passages.

Bilingual target children at each grade level, except kindergarten, were administered the reading achievement battery in both English and Spanish in March/April of each year. The monolingual children were tested with only the version of IRAS appropriate to their language group.

Standardized Achievement Tests. Standardized achievement tests in English were administered by the Texas schools in the Spring of each year to all students, starting with Grade 1. The three border area sites used the California Achievement Test. The Iowa Test of Basic Skills was administered in the urban site. The Central Texas Area--Site 3 used the Comprehensive Test of Basic Skills. These data were obtained from the schools each year for all target students.

Standardized achievement tests in Spanish were not administered systematically, nor to any great extent, by any of the schools in the study. Similarly, no such tests were administered in the Northern Mexico site.

Cognitive style. As noted in the section above on the selection of target students, all students in each of the target classrooms, at the time of entry into the study, were administered two cognitive measures: the Children's Embedded Figures Test and the Matching Familiar Figures Test. These tests were then readministered to the target children only in the fall of each year that they remained in the study. Description of these tests is provided in the section referred to above (see page 21).

Cognitive development. The Cartoon Conservation Scales (De Avila, 1976) is an individually-administered, Piagetian-based measure of cognitive development. It was developed as a means for assessing the intellectual development of children in a manner that is fair to children of diverse linguistic and ethnic backgrounds. For this reason, it was selected as a measure of cognitive development for the purposes of the study.

The instrument consists of five subscales for each of two levels. Level I, for use in grades K-3, assesses conservation of number, substance, distance, identity, length, and egocentricity. Level II, designed for use in grades 4-6, contains measures of conservation of substance and distance, horizontality of water, class inclusion, probability, and egocentricity.

Only the target students in the study were tested with the Cartoon Conservation Scales. The testing was carried out yearly during the months of November through January.

Instruction and Teacher Characteristics

Instruction. A coordinated system of teacher interviews, teacher checklists, and monthly classroom observations provided rich and extensive data for the purpose of documenting and describing in detail the instructional program each child experienced over the course of the study. A brief overview of these measures is presented below.

Teacher interviews. Each of the target teachers was interviewed twice during the school year, in the early fall and in late spring. In the site selection process much was learned about the nature of the school program, the sample classrooms, the student population, and the background and training of the teachers and teacher aides. The fall interview was used to clarify, on an individual basis, the following: (a) organization of students for reading; (b) basis for grouping; and (c) schedule for reading activities. This interview was also used to establish rapport between the teachers and the research staff and to orientate the teacher to the study in general and to the scheduled activities for the year. In the spring interview, the staff obtained feedback from the teachers on the instruments and procedures used in the study.

Inventory of Bilingual Instruction. This instrument is a questionnaire designed to elicit information from the classroom teacher that allows each classroom program to be defined in terms of variation on three major components:

1. percent of instruction time for language arts devoted to language arts in Spanish;

2. percent of instruction time for content areas other than language arts taught in Spanish; and

3. grade levels at which such instruction is provided.

Essentially, the teachers are asked for their current and projected daily schedule of classroom activities. For each activity indicated, information is sought regarding the language categories of the students within each instructional activity, the primary instructor, the language of instruction, and the language of the materials. This instrument was administered to all target teachers in the fall of each school year.

Reading Checklist (Teacher Instructional Plan). To supplement and verify the representativeness of the information obtained in the classroom observations (discussed below), the teachers in the study were interviewed three times during the school year by SEDL staff using the Reading Checklist. Essentially, the Checklist obtains from the teachers, in computer-compatible codes, their instructional plans for reading over a two-week period for each of the target children or groups of children in their classes. For each group of children, the strategies or skills that were taught during the preceding two-week period are listed in order of emphasis. For each of these listed, the following are indicated: instructional focus, material (type, title, and section of the book), type of activity, language of instruction, instructor (e.g., teacher or teacher aide), role of the instructor, and total minutes devoted to the teaching of the strategy/skill over the two-week period.

Classroom Observation. Observations, using an adapted version of the Reading and Mathematics Observation System-RAMOS (Calfee & Calfee, 1976) was carried out in each target classroom once a month from October through April/May. The purpose of these observations was to collect information on certain instructional variables, teacher behaviors, and student response to the instruction. Instructional variables of interest included the following: frequency and duration of instructional events that relate to a particular approach to the teaching of reading, role context of particular teacher behaviors, sequencing of instructional units, patterns of language use (Spanish and English) during reading instruction, general structure and organization for reading instruction, and instructional roles of the teacher and teacher aides.

The RAMOS system provides real-time documentation of classroom instruction, unlike the time-sampling methods used in many other systems. Any time there is a change in skill or activity in one or more of the categories of instruction, this change is identified and recorded. Such documentation allows for an assessment of the time spent by teachers and students in each of the several categories of RAMOS. The system is used by a trained observer who records information on an Event Form using computer-compatible mnemonic codes. These codes are then translated into a detailed account of the instructional process during a given period of reading instruction. Repeated observations throughout the year, as well as across years, yields an extensive documentation of the instruction each target child has experienced.

Teacher characteristics. Two types of measures were used to obtain information about the characteristics of the teachers in the study. One obtained information about the professional training, experience, and language skills of the teachers; the other assessed the two dimensions of cognitive style selected for use with the children in the study.

Teachers' background and language skills. Information about the teachers' professional training, experience, and language skills was obtained through the Survey of Teachers' Background and Language Skills (SEDL, 1979). This is self-report instrument which was completed once by each target teacher during the teacher's initial year of participation in the study.

The instrument consists of three sets of items. One elicits information relative to the following: age, sex, origin of descent, highest degree earned, present and past teaching assignments, and certification status. Another set of items inquires about the teacher's language background and abilities. These items provide information about the childhood language of the teacher, other languages currently spoken, and the situation in which each of these languages was learned. Teachers who have skill in Spanish are asked to indicate their perceived ability to perform in particular language situations that involve different levels of speaking, reading, and writing Spanish. The final set of items provides a summary of courses, typically offered to prepare teachers to teach students of limited English-speaking ability, that the teacher has taken both at the college/university level and in local staff development programs.

Cognitive style. The adult version of the two measures of cognitive style used with the students were administered once to the target teachers during each teacher's initial year of participation in the study. The Group Embedded Figures Test-GEFT (Consulting Psychologists Press, 1971) is a measure of field independence/field dependence. The test consists of a set of complex geometric patterns in which simple figures are embedded. The score is the total number of simple forms correctly identified within a set time limit.

The adult version of the Matching Familiar Figures Test (Kagan, 1966), while quite similar to the child's version, utilizes different standard stimuli and variants. It also differs slightly in format and in number of items. The reflectivity/impulsivity classification is based on the time it takes the subject to indicate the first response (latency) and the total number of incorrect responses made on each item.

Data Collectors

All data that required direct teacher input was collected by full-time members of the SEDL research team. This required systematic visits to the research sites. Formal classroom observations and the collection of student data were carried out by a data collection team from each of the sites. The team consisted, in most cases, of two people who were not employed in other endeavors. All met the following criteria: residents of the local community, experienced teachers, Hispanic and fluent speakers of English and Spanish, and acceptable to the school district. In all cases, the school district administration provided a list of acceptable and available people who were then screened by the SEDL staff.

In the late summer of the year of initial entry into the study, extensive training was provided by the SEDL research staff for the data collection team. Generally, the first training session, of four to five days in duration, concentrated on the instruments and data collection procedures to be used in the first two months of school. A second session was held in early October at which time training on the next set of instruments was provided and additional training and clarification was carried out for the observation system. A third training session was held in the spring which focused on the administration of the reading achievement instrument and the procedure for the collection of standardized test scores and attendance data. Telephone and mail communication was frequent between the SEDL staff and the local data collectors throughout the year. In subsequent years, training sessions were held as needed, typically two per year. On one occasion, after all cohorts were in the study, all of the data collectors were brought to SEDL for two days of training in an effort to ensure uniformity of test administration and collection procedures across the sites. The study was fortunate in acquiring the services of exceptionally capable people at the sites who, for the most part, remained with the study for its duration.

References

Brown, A. L., Campione, J. E., & Day, J. D. (1981). Learning to learn: On training students to learn from texts in the intermediate grades. Educational Researcher, 10, 14-25.

Calfee, R. C. (1977). Assessment of independent reading skills: Basic research and practical applications. In A. S. Reber and D. L. Scarborough (Eds.), Toward a psychology of reading. Hillsdale, NJ: Lawrence Erlbaum, Associates.

Calfee, R. C. (1981). Cognitive psychology and educational practice. In D. C. Berliner (Ed.), Review of research in education. Washington, DC: American Educational Research Association.

Calfee, R. C. (1982). Cognitive models of reading: Implications for assessment and treatment of reading disability. In R. N. Malatesha and P. G. Aaron (Eds.), Reading disorders: Variety and treatments. New York, NY: Academic Press.

Calfee, R. C., & Calfee, K. H. (1976). Reading and mathematics observation system-RAMOS/II (rev.). Unpublished manuscript, Stanford University, Stanford, CA.

Calfee, R. C., & Associates. (1977). Stanford foundation skills test (group form, rev.). Unpublished manuscript, Stanford University, Stanford, CA.

Calfee, R. C., & Peña, S. C. (1978). Stanford foundation skills test. Stanford University, Stanford, CA.

Calfee, R. C., & Calfee, K. H. (1979). Interactive reading assessment system (IRAS). Stanford University, Stanford, CA.

Calfee, R. C., & Calfee, K. H. (1981). Interactive reading assessment system (IRAS). Unpublished manuscript (rev.), Stanford University, Stanford, CA.

Calfee, R. C., & Shefelbine, J. L. (1981). A structural model of teaching. In A. Lewy and D. Nevo (Eds.), Evaluation roles in education. New York, NY: Gordon & Breach.

Calfee, R. C., & Sutter, L. (1982). Oral language assessment through formal discussion. Journal of Language Disorders, 2, 45-55.

Calfee, R. C., Calfee, K. H., & Peña, S. C. (1979). Interactive reading assessment system-Spanish. Stanford University, Stanford, CA.

Consulting Psychologists Press, Inc. (197). Group embedded figures test. Palo Alto, CA.

Cronbach, L. J. with the assistance of Deken, J. K., & Webb, I. V. (1976). Research on classrooms and schools: Formulation of

70

questions, design, and analysis. (Stanford Evaluation Consortium Occassional Papers). Stanford, CA: Stanford University, School of Education.

Cummins, J. (1980). The cross-lingual dimension of language proficiency: Implications for bilingual education and the optimal age question. TESOL Quarterly, 14, 175-187.

De Avila, E. A. (1976). Cartoon conservation scales. Corte Madera, CA: Linguametrics Group, Inc.

De Avila, E. A., & Duncan, S. (1979). Relative linguistic proficiency and field dependence/independence: Some findings on linguistic hetrogeniety and cognitive style of bilingual children. Austin, TX: Southwest Educational Development Laboratory.

Duncan, S. E., & De Avila, E. A. (1976). Student operational language assessment scale. Austin, TX: Southwest Educational Development Laboratory.

Goody, J., & Watt, I. (1963). The consequences of literacy. Comparative Studies in Society and History, 5, 304-345.

Gough, P. B., & Hillinger, M. L. (1980). Learning to read: An unnatural act. Bulletin of the Orton Society, 30, 179-196.

Guszak, F. J. (1972). Diagnostic reading instruction in the elementary school. New York: Harper & Row.

Heath, S. B. (in press). Ways with words: Ethnography of communication in communities and classrooms.

Kagan, J., Rosman, B., Day, D., Albert, J., & Phillips, W. (1964). Information processing in the child. Psychological Monographs, 78, 1-37.

Kagan, J. (1965). Reflection-impulsivity and reading ability in primary grade children. Child Development, 36, 609-628.

Kagan, J. (1966). Reflection-impulsivity: The generality and dynamics of conceptual tempo. Journal of Abnormal Psychology, 71, 17-24.

Lesiak, J. (1970). The relationship of the reflection impulsivity dimension and the reading ability of elementary school children at two grade levels. Unpublished doctoral dissertation, Ohio State University.

Mace-Matluck, B. J., & Domínguez, D. (1978a). Informal reading inventory: Keys to reading. Austin, TX: Southwest Educational Development Laboratory.

Mace-Matluck, B. J., & Domínguez, D. (1978b). Inventario informal de la lectura. Austin, TX: Southwest Educational Development Laboratory.

Mace-Matluck, B. J., Domínguez, D., & Padilla-Hajjar, R. (1978). Inventario informal de la lectura. Austin, TX: Southwest Educational Development Laboratory.

Mace-Matluck, B. J., Tunmer, W. E., & Domínguez, D. (1978). Ethnographic verification of language status. Austin, TX: Southwest Educational Development Laboratory.

Mace-Matluck, B. J., Tunmer, W. E., & Domínguez, D. (1979). Oral language proficiency rating scale. Austin, TX: Southwest Educational Development Laboratory.

Olson, D. R. (1977). From utterance to text: The basis of language in speech and writing. Harvard Educational Review, 47, 257-281.

Olson, D. R. (1980). On the language and authority of textbooks. Journal of Communication, 30, 186-196.

Simon, H. A. (1978). Information-processing theory of human problem solving. In W. K. Estes (Ed.), Handbook of learning and cognitive processes: Human information processing, Vol. 5. Hillsdale, NJ: Lawrence Erlbaum, Associates.

Simon, H. A. (1981). The sciences of the artificial (2nd ed.). Cambridge, MA: MIT Press.

Southwest Educational Development Laboratory. (1979). Survey of teacher's background and language skills. Austin, TX: Author.

Sternberg, R. J. (1967). Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities. Hillsdale, NJ: Lawrence Erlbaum, Associates.

Stone, M. K. (1976). Correlates of teacher and student cognitive style. Beginning teacher evaluation study, Phase IV, 1974-75. (ERIC Document Reproduction Service No. ED 131 120)

Texas Education Agency. (1976). Annual statistical report (1973-1974): Part I. Austin, TX: Texas Education Agency.

Texas Education Agency. (1978). State plan for bilingual education: A system for meeting the needs of limited English-speaking students in Texas. Austin, TX: Texas Education Agency.

Texas Education Agency. (1980). 1979 Fall survey. Austin, TX: Texas Education Agency.

Witkin, H. A., Oltman, P. K., Raskin, E., & Karp, S. A. (1971). A manual for the embedded figures tests. Palo Alto, CA: Consulting Psychologists Press, Inc.

Zamora, J. E. (1977). A status survey of Texas bilingual-bicultural education programs. Unpublished doctoral dissertation, The University of Texas at Austin.

APPENDIX A

Site Descriptions

SEDL BILINGUAL READING STUDY

Summary Description of Project Sites

Texas Border Area - Site 0

District Setting

The district serves a 400 square mile area in a rural section of south central Texas. It spreads out over the eastern part of Starr County and serves three population centers. The area served by the school district is in close proximity to the US-Mexico border and is situated approximately halfway between Laredo and Brownsville. The general characteristics of the school district, as of 1982, are listed below:

Size: Schools - 8 (1 high school, 9-12; 2 junior high schools, 6-8; 2 intermediate schools, K-5; 3 elementary schools, K-4).

Students - 5,460

Professional staff - 322 (of which 281 are classroom teachers).

Paraprofessional staff - 310 (teacher aides, secretaries, and assistant nurses).

SES (community): approximately 88% of the families represent low income households (per capita income in Starr County is $2,668 a year).

Ethnic Composition:

Elementary	Secondary
Hispanic - 97.0%	Hispanic - 99.0%
Anglo - 03.0%	Anglo - 01.0%

Distribution of Hispanic Population: The entire school population is essentially Hispanic.

Level of Support - ADA: 4,765

Per Pupil Expenditure: $2,427.66

Enrollment Trends: Increasing enrollment at the rate of 2.5% to 3.5% yearly.

The school district is located in southcentral Texas in close proximity to the US-Mexico border. It serves a large rural area, about the size of the state of Rhode Island, that lies in the eastern part of Starr County.

Starr County, the 35th poorest in the nation, has a population of 18,000 people, 98% of which are of Mexican descent. Spanish families settled in Starr County as early as 1765. From 1765 to 1848 the Hispanic population flourished, and the Spanish language, customs, and traditions predominated.

After 1848, with the signing of the Treaty of Guadalupe Hidalgo, Anglos began coming into the region. They were mostly men who eventually intermarried with the Hispanics in the area. The number of Anglos increased in the area with the coming of army personnel to the local calvary post in the early 1900s. These soldiers married girls from the immediate community, thus bringing Anglo names into the community. As time went by, the parents died or left the community; the off-springs often married spouses from the area and remained in the community. Today there are many families in the region with English family names, but they are ethnically Hispanic and have integrated into the Spanish-speaking community. In later years, very few Anglo or English-speaking families moved into the area. On the other hand, a large number of Mexican families have crossed the border and settled in the region.

The socioeconomic level of the people living within the region served by the school district is one of the lowest in the nation. Except for a narrow strip of land stretching along the banks of the Rio Grande River (approximately one to two miles wide and 15 miles long) the land within the district is not productive agricultural land. Most is hilly brush land with a small acreage fit for dry land cultivation. Next to farming in irrigated lands, ranching is perhaps the most productive. The frequent lack of rain curtails full production, however. Other than the oil and gas industry in the region (and this has slowed to a very small amount) few industries have come into the area.

Unemployment runs high throughout most of the year; the number of unemployed people at any given time usually ranges from 15% to 50%. Due to limited employment opportunities, a large segment of the population migrates yearly to work in agricultural crops along the "migrant trail" from Texas to the state of Washington. This pattern has been engendered for generations, thus contributing generally to a low level of education for the people in the region. Nonetheless, the district's school officials point with pride at a cadre of their graduates over the past decade or so who have let poverty be, rather than a stigma, an incentive for high achievement and escape to a better life. These include at least 10 medical doctors, nine attorneys, seven pharmacists, four dentists, seven high-ranking school administrators, five engineers, an architect, a research scientist, a millionaire grocery chain operator, and innumerable teachers. School officials also point with pride to the fact that, in spite of a great many deterrents related to poverty, their students value education and some 67% of the 1982-83 high school graduating class planned to further their education.

The main administrative complex of the school district is located in what used to be the cavalry post. The district has gradually turned the historical army post into one of the most beautiful and functional school complexes in South Texas. It is a unique complex in which one can readily enjoy the interesting configuration of the parade grounds and the colonial Spanish architecture of the buildings which now have been modernized and house the district's intermediate school. The senior high school, one of the district's junior high schools, the district's auditorium, and the central administrative offices (all of which are modern structures) are also located within the complex.

The district serves approximately 5,460 students in kindergarten through grade 12 in eight campuses, five of which are situated within the city limits of the largest of the population centers served by the district. The remaining three schools (one junior high school and two elementary) serve two rural communities, the most distant lying some 13 miles from the central administrative offices.

The school enrollment over the past five years has been increasing.This increase, which is estimated to be from 2.5% to 3.5% yearly, has come about by an influx of families from Mexico into the region. Previously, farm workers would leave their families in Mexico and come to work during the agricultural season in the United States, returning to their homes when their work was completed. However, in the last ten years they have brought their families with them and have settled in the communities along the border. This accounts for much of the increase in the school population over the past several years.

Bilingual Education Program

In 1972 the school district implemented through ESEA Title VII funds a bilingual-bicultural curriculum in grades kindergarten through grade 3 and in two sections of grade 4. In subsequent years this program was expanded to provide special language assistance instruction to children in kindergarten through grade 10. Other funding agencies became the primary sources for these classroom programs. A State Bilingual Program was implemented to serve children in kindergarten through grade 5; a program of ESL instruction was provided through ESEA Title I Migrant funds for children in grades 6-10. At present these two programs serve some 3,354 students (61% of the student population). Of these, approximately 35% are served by the bilingual program with the remaining 27% receiving ESL instruction through the Title I Migrant program. Some 86 teachers (64 in K-5; 22 in grades 6-10) provide the basic instruction in these programs.

The primary goal of the district's special language assistance programs, in addition to promoting academic achievement, is the development of self-confidence and self-esteem as an approach to reducing the dropout rate and keeping the students in school sufficiently long to raise the level of education of the current generation of students. The curriculum plan for the bilingual program specifies instruction in the following areas: oral language development, mathematical concepts using the ESL approach, English-as-a-Second Language, Science in both languages with an ESL approach provided in the English component, Art and Music, Reading and Writing in English and Spanish, and cultural heritage. Depending on the school and grade level(s) involved, instruction in Spanish reading and communication skills is provided on a daily basis (usually 20 minutes) or for a longer period on certain days of the week. The building principal, under the direction of the Office of the Assistant Superintendent for Curriculum, is responsible for the administration and supervision of all instructional programs in her/his building.

School Sites

The target students from this site were drawn from two elementary schools: School A and School B. The initial cohort consisted of 10 kindergarten students and 10 first graders from School A and a similar sample from School B. The following year an additional cohort of students entered the study: 10 kindergarten students from School and 10 from School B. In all, 60 children from this site participated in the study.

School A. School A serves grades K-4 and is located near the center of town in close proximity to the main school complex. The school's population of some 717 students is essentially Hispanic (96%) from low income households (some 86% are so classified).

The classes are housed in a colonial Spanish style school building built some years ago and two or three (depending upon the year) temporary buildings adjacent to the main building. During the course of the study a new wing was added to the main building, and renovation of the main building and one of the temporary buildings occurred. During the period of the construction the fourth grade classes were housed in the intermediate school in the main complex (School C, discussed below).

The staff at School A consists of the principal, the assistant principal (added to the staff during the final year of the study), 33 teachers, 28 teacher aides, and two secretaries. Instructional leadership at the building level is provided by the principal and the assistant principal. The principal at the school when the study was first implemented was transferred to a neighboring school at the end of the second year, thus during the four years of the study, two people served in the role of principal.

Of the 33 teachers at the school most are fluent Spanish speakers who have grown up in the region. Generally, they live in the community or its environs, although a few commute to work from other small towns or communities in the area. There is a moderate turnover of teachers yearly due in part to retirement, to transfer to other schools in the district, or to younger teachers moving to the larger population centers. Almost without exception, the teacner aides are Hispanics who reside in the community.

The classrooms at each grade level were stocked with essentially the same basic curriculum materials, which in general appeared to be adequate with the possible exception of Spanish reading and language arts materials in the early years of the study. The curriculum materials remained relatively stable throughout the study, as did program practices, with the exception of the following: a new basal reading series in English was adopted and implemented during the 1980-1981 school year at which time the Wisconsin Design curriculum materials were discontinued; a Spanish basal reading series was adopted in 1982 which replaced "public domain" materials that were developed some ten years earlier by the regional service center and by the regional laboratory.

Class size ranged from approximately 25-30 in kindergarten and grade one and from 30-38 in grades 2-4. The classes were self contained in K-3; the classroom teacher taught all subjects with the exception of physical education which was provided by the P.E. teacher. At fourth grade, the instructional program was departmentalized, with students changing classes approximately every 45-50 minutes.

School B. Located some 13 miles out in the countryside, School B serves some 363 children in K-5 who reside in a small community near the school or in the rural area stretching out several miles from the school. Many of the children are bussed to school. Most are Hispanic, Spanish-speaking children (99%) from low income households (88%). Many of the parents of these children work in agriculture, and a good number of them are migrant workers.

The school building, a low, modern structure in good repair, is situated adjacent to the junior high school building, also a modern structure, which serves children from that attendance area. Enrollment at School B necessitates the use of a long, temporary structure adjacent to the school which houses three-to-four classes and some of the special programs (e.g., computer lab).

The school staff consists of a principal, 20 teachers, 22 teacher aides, and one secretary. The principal serves as the instructional leader at the building level. Midway through the study the principal at School B was transferred to School C and was replaced at School B by an experienced principal from one of the others schools in the district. Teacher turnover is somewhat high at this school, although there are a few teachers there who were there some years prior to the study and remained there throughout the course of the study. The teacher aides, for the most part, live within the attendance area of the school, as do some of the teachers. Some teachers, however, live in the largest of the district's three population centers or other communities and commute to work.

Curriculum materials and program practices were essentially the same as those in School A, although the organizational structure differed somewhat. Classes at all grade levels were self-contained; the classroom teacher with the assistance of a teacher's aide, provided all of the instruction for her/his students except physical education and special program instruction such as computer-assisted instruction in math and supplemental reading and language arts instruction which was provided in pullout classes.

Class size ranged from approximately 20-25 in kindergarten and grade one and increased to 25-30 at other grade levels. Students of parents who were migrant workers usually withdrew from school in early to mid-April and usually returned to school the following year at the beginning or within the first weeks of school. A few of these children, however, did not re-enroll in the school until well into November as was generally the case some years ago.

School C. As indicated above, the children in the study from School A attended School C during their fourth grade year during the construction of the new wing at School A. School C, an intermediate school that usually serves grades 1-5, is housed in the renovated living quarters of the former army post. During the year that the fourth grade students were there, they formed a part of the school population, typically some 704 students of which approximately 96% are Hispanic from low income households (88%). Classroom practices and curriculum materials for these children were similar to those of School A described above.

The Reading Program

The location of the school district and the language usage patterns in the region virtually ensures that the overwhelming majority of the children enter school as fluent speakers of Spanish; some are fluent speakers of English as well while others have a working knowledge of English, and yet others have no or only limited knowledge of English.

In the early years of the study, 1978-1980, there was a strong emphasis on the teaching of reading in English for all students, regardless of the student's oral English proficiency at the time. Teachers frequently stated that many of the children had limited exposure to English outside of the school, and they expressed the fear that unless the school concentrated on English many of the children would not learn English sufficiently well to function in an English-medium society. The reading program during those years reflected those concerns.

The configuration of the reading program during the early years of the study was as follows. All children began reading instruction in both languages. Usually English reading was taught in the morning. The children were grouped within the class (usually 3-4 groups) for instruction on the basis of reading skill in English. The Wisconsin Design Management System served as a framework of skills to be taught, and basal readers (The Economy Series) were used as the basic instructional materials. Each group received approximately 20 minutes of direct instruction provided by the classroom teacher daily; the remainder of the period (45 minutes to one hour) was spent in independent work, usually based on workbooks, ditto sheets, or chalkboard assignments. Later in the morning or in the early afternoon a period of time (usually 20 minutes) was scheduled fr Spanish reading instruction. By comparison to the English reading ma ials, the Spanish reading materials were somewhat sparse. Materials used were the BOLaR Series and the Spanish Roll. Teachers made use of the chalkboard and charts to present word recognition skills to the whole group. This activity was usually by oral and/or silent reading of a limited amount of assigned material followed by a teacher-led discussion about the material read.

In keeping with the state schedule of textbook adoption, a new textbook adoption for reading materials occurred for the beginning of the 1981-1982 school year. The Macmillan series was adopted by the district for the English reading component, and the Economy Series was selected for the Spanish reading component. At this point the practice was implemented in which formal reading instruction in Spanish was provided only for those children who were determined to be Limited English Proficient. These children were grouped for instruction according to reading ability and received their basic reading instruction in Spanish while the other children in the class who were English Proficient received their basic instruction in English, as described above. At fourth grade, however, instruction at School A is departmentalized. Children are scheduled, by classroom, for one hour a day with the reading teacher and receive their basic reading instruction during that period.

Transfer Criteria:

As children in the Spanish reading component of the program attain specific skill in reading in Spanish and in oral English development, they are gradually phased into the reading program in English. The school district has recently specified those skills and abilities that are to be considered when transfering students, and teachers have been informed of the district's policies and procedures in this regard.

District-Wide Staff - Reading Program:

In addition to Assistant Superintendent for Curriculum and the Administrative Assistant in Curriculum, the district provides a Supervisor of Communication Skills. This person works directly under the supervision of the Office of Assistant Superintendent for Curriculum and assists in the dissemination of information to the school staff and is responsible for assisting teachers on a day-to-day basis in the implementation of the school district's adopted program in reading.

Staff Development

In 1979 a Five-Year Educational Improvement Plan, based on the findings of a self-study conducted in 1977-1978, was implemented. The priority student needs areas identified were reading and computational skills. The Five-Year Plan specified a program of staff development. A minimum of two workshops (usually a day each) were planned and carried out each year which addressed identified needs in reading. Topics for these workshops included the development of: expressive language, work attack skills, interpretive and critical comprehension skills, writing skills, study skills, and reading skills in the content areas. In addition, two workshops were planned and carried out each year which focused on promoting teacher competencies in curriculum development in reading.

80

Some Problems and Concerns

The problems and concerns most frequently expressed by the school staff during the course of the study were the following:

1. The lack of exposure to English outside of the school. Many of the children reside in rural, isolated areas and are bussed to school. When they leave school at the end of the day, they return to an all-Spanish-speaking environment. Even the children who reside in the population centers may have limited exposure to English since Spanish can, and often does, serve as the primary vehicle of communication in all domains of life and thought.

2. Problems related to poverty in the region. These include higher incidence of disease, substandard housing, malnutrition, and unemployment than in most parts of the country.

3. Lack of access to educational and cultural enrichment centers such as those found in metropolitan centers, such as public libraries, museums, theaters, zoos, and concerts.

4. Lack of literacy materials in the home environment. Many of the students come from poor families and do not grow up with accessability to books, magazines, and newspapers.

5. Difficulty in attracting and keeping well-prepared teachers in the district. Even though the salary schedule for the district includes increments above the state-based pay schedule, the location and character of the region makes it difficult to attract new teachers from outside the region, in particular, to the school district.

6. Low performance of the students on standardized achievement test scores. While progress has been made in student performance on these tests, scores are still uniformily low in reading, language, and math when compared to the national average.

SEDL BILINGUAL READING STUDY

Summary Description of Project Sites

Texas Border Area - Site 1

District Setting

The district is located in a rural area in close proximity to the US-Mexico border and is situated in far west Texas, 21 miles east of El Paso. The general characteristics of the school district, as of 1982, are listed below:

Size: Schools - 2 (1 elementary, Pre-K through 6; 1 secondary, 7-12)

Students - 798

Teachers - 40

SES (community): low income households (medium income in 1979 - $9,639)

Ethnic Composition:

Elementary		Secondary	
Hispanic	97.9%	Hispanic	95.9%
Anglo	01.8%	Anglo	03.1%
Black	00.2%	Black	00.9%

Distribution of Hispanic Population: The entire school population is essentially Hispanic.

Level of Support - ADA: 666

Per Pupil Expenditure: $1,457

Enrollment Trends: Increasing enrollment over the past six years (34.4%);

Projected enrollment trends - overall increase of 24.7% in next four years (1986-1987); overall increase of 43.4% over next 10 years (1991-1992).

82

The school district serves grade levels pre-kindergarten through grade 12. From a declining enrollment in the middle 1970s, the district has been experiencing a sharp increase in enrollment over the past six years (34.4%), with an expected overall increase in the next ten years of 43.4%. The school population is made up almost entirely of Spanish-speaking students from low income households. The history and culture of this community dates back some 400 years. Many of the families have been in the area for generations. For the most part, they are not migrants, but live and work in and around the community. New people presently coming into the community are previous residents of Mexico. The school population of 798 students is served by one elementary school and one secondary school.

Bilingual Education Program

The district provides a bilingual education program (grades K-12) to serve limited English proficient (LEP) students. The primary goal of the bilingual program is to prepare LEP students to be able to function in an all-English instructional program. To meet this goal, the student is instructed in his native language while at the same time receiving English-as-a-Second Language (ESL) instruction. The components of the program are: Reading in Spanish (K-6); ESL (K-12); and Math, Science, and Social Studies in English with concept development in Spanish (K-12). The principal administers the bilingual program, however, during two years of the four years of the study a bilingual programs coordinator was employed, with Title VII funds, to assist in planning and carrying out staff development activities, matters of curriculum, testing, and in diagnosis and placement of students.

School Site

The target students from this site were drawn from the district's single elementary school which houses pre-kindergarten through sixth grade. The school population of approximately 480 students is essentially Hispanic (97.9%) of Mexican origin. The school serves a rural community made up primarily of agricultural and blue collar workers, most of whom represent low income households. The classes are housed in a Spanish colonial style school building built in the 1920s and in four temporary buildings adjacent to the main building. The building is in need of renovation and repair; during the final year of the study, the district was able to get bonds approved to renovate the building and to add a modern wing to accommodate their increasing school population.

Administrative/Supervisory Personnel. The principal serves as the instructional leader at the building level. During the course of the study (four years), there were major changes in the administrative/supervisory personnel. At the end of the second year, a new central office staff was hired (superintendent, curriculum director), the elementary principal moved to the high school principal position, and one of the teaching staff was hired as principal of the elementary school. During the final year of the study, the former elementary principal returned to his original position. For the two middle years of the study, a coordinator for the bilingual program was employed with Title VII funds. There was a change in personnel for this position each year.

Teaching Staff. Approximately 20 teachers are employed at the elementary level. There is a high rate of teacher turnover from year to year. Few of the teachers live in the community; most commute either from El Paso or from larger neighboring towns. Approximately one-half of the teachers are Hispanic and fluent speakers of Spanish. Depending on funding, a small number of aides are employed; these generally are drawn from the community.

Classrooms. The classrooms at each grade level were stocked with essentially the same basic curriculum materials, which in general appeared to be adequate. The curriculum materials remained relatively stable throughout the course of the study, as did program practices, with the exception of the following: a new basal reading series in English was adopted and implemented during the 1980-81 school year; an Uninterrupted Sustained Silent Reading Program, a supplement to the reading program, was implemented during the final year of the study. This consisted of a 15-minute period (from 10:00 am-10:15 pm) daily in which all activities stopped, and children and adults alike took out reading materials of their choice and read uninterrupted for 15 minutes.

Class size ranged from approximately 20-25 students per class in kindergarten and grade one (during the first year of the study) and increased to 30-35 students per class at all grade levels in the later years of the study. All of the teachers assigned to the Spanish reading program held the State of Texas Bilingual Endorsement Certification. The training and teaching experience of the teaching staff varied from year to year. Students changed instructional environments frequently during the course of a day due to the following: departmentalization at grades 4-6 and for certain subjects for all students (e.g., P.E., Music); ability grouping procedure for reading and math for all students except at the kindergarten level; and pull-out instruction for Migrant and Title I programs.

The Reading Program

The district provides an Interdisciplinary Basic Skills in Reading Program for all students. Monolingual English-speaking and English proficient bilingual students receive reading instruction in English only. Limited English-speaking students receive initial reading instruction in Spanish. All students are grouped according to their instructional level in Reading, Language Arts, and Mathematics. Students are assigned to a teacher for reading instruction on the basis of grade and language status. For example, within a first grade class of fluent English-speaking students, the teacher groups the students on the basis of reading ability, usually resulting in 3-4 groups. Each group receives approximately 15-20 minutes of direct instruction provided by the teacher daily. The remaining portion of the period (usually one hour) is spent in independent work, usually based on workbooks, ditto sheets, or chalkboard assignments. Classes for Spanish-dominant LEP students are similar in their organization and content. At the primary grades, the homeroom teacher is responsible for instruction in all other subject areas except P.E. and Music, which is usually taught by a specialist. At the upper grades, instruction is departmentalized; certain teachers teach only one subject (e.g., science, social studies, music, P.E.), and students are scheduled through those classes. Teachers in the science and social studies areas also reinforce reading skills in their subject area classes.

Basal reading series provide the foundation for both the English and Spanish reading programs. In these materials, skills and vocabulary are sequenced and increase in difficulty from one level to the other. In keeping with the state schedule of textbook adoption, a new textbook adoption for reading materials occurred at the beginning of the third year of the study. The American Book Company series was adopted for the English reading component and a new version of the Santillana Publishing Company materials replaced the older version of that series which was used in previous years. Previous reading texts for English were the Harcourt Brace series. The basal reading program was reinforced by the Uninterrupted Sustained Silent Reading Program, described above, during the last year of the study.

When students in the Spanish reading component of the program reach "transfer criteria," they are placed into a transition English reading program (Santillana materials). When the transition program is completed, the student is evaluated, and if she/he meets the criteria, is assigned to the all-English, mainstream instructional program.

Transfer Criteria:

Pre-Transfer Stage

- Score 95-100 points on Rayuela (Libro 1) End of Book unit test
- Score a Level 4 or above on the English LAS I or II
- Reads at grade level as measured by an IRI in Spanish (Rayuela)

Transfer Reading Stage

- Score 100 points on Able to Read (Santillana transfer materials) End of Book unit test

Criteria for Exiting the Program

- Score 100 points on Lickety-Split (Santillana transition materials) End of Book unit test
- Score above 40% on CAT

A two-year tracking system has been designed and implemented in an effort to monitor students' progress after exiting the program. If a student's grades and/or scores on the California Achievement Test (CAT) drop during this two-year tracking period, the student is reclassified as a LEP and again placed in the bilingual program for further help in the identified need areas.

Staff Development

The school district typically provides inservice training during five days prior to the opening of school and periodically during the school year. In the 1982-83 school year, the teachers received seven days of inservice, five of which were just prior to the opening of school. Topics treated included curriculum guides, successful practices, materials available in the district, special education, discipline, ESL, teacher effectiveness training for bilingual classrooms, and orientation to the California Achievement Test.

Some Problems and Concerns

During the course of the study, the superintendent and principal expressed concern about the lack of student exposure to the "outside world." Students and their families occasionally go into El Paso for business or recreational trips, but they have close ties across the border and seldom leave the area to seek education or job opportunities. They worried about the social and economic mobility of their students, and have attempted, within the limitations of their resources, to provide activities such as field trips to expand their students' knowledge of the region and of career opportunities.

The increasing population to be served, in the face of limited physical facilities and revenue sources, presents a real problem for the district, as does the recruitment of well-trained teachers in the face of a high teacher turnover each year.

Other concerns expressed were related to language assessment and low performance of the students on standardized achievement tests. The Language Assessment Scales was used in each year of the study. However, the district is currently seeking help from the Regional Service Center and the local universities in designing better ways of evaluating their students' language ability. To address student performance on standardized achievement tests, the district developed and implemented in 1980 a five-year plan for improving student achievement. They identified needs and set priorities in each of the curriculum areas. The focus of the five-year plan is program improvement in Reading, Math, and Career Education.

SEDL BILINGUAL READING STUDY

Summary Description of Project Sites

Texas Border Area - Site 2

District Setting

The district is located in an unincorporated town with an estimated population of 4,213 in the general area. The area lies 30 miles east of El Paso and is situated approximately 15 miles north of the US-Mexico border. The general characteristics of the school district, as of 1982, are listed below:

Size: Schools - 4 (1 primary school, K-3; 1 elementary school, grades 4-6; 1 junior high school, grades 7-8; 1 high school, grades 9-12).

Students - 1,625

Teachers - 103

SES (community): low to lower middle income households

Ethnic Composition: 90% Hispanic

Distribution of Hispanic Population: The entire school population is essentially Hispanic

Level of Support - ADA: 1,544

Per Pupil Expenditure: $1,964

Enrollment Trends: Increasing enrollment at the rate of approximately 4% over the past few years.

The school district serves approximately 1,625 students (grades K-12) in four schools: the primary school (K-3); the elementary school (4-6); the junior high school (7-8); the high school (9-12). All four schools are located near the center of town in close proximity to each other. The district serves a 71 square-mile area, about one-half of it lying in rich irrigated valley farmland and one-half in barren sand hills. Many of the students are transported to and from school over four established bus routes. The students, generally, are from low to lower middle income households. The parents of some work either in agriculture in the area, in local businesses, or at the small manufacturing plant in the area; others commute to El Paso to work in businesses and industry there. The student enrollment has been relatively stable, showing a 4% increase over the past few years.

The Bilingual Program

Bilingual education is provided through fifth grade; English-as-a-Second-Language (ESL) is offered at all grade levels (K-12). During the first two years of the study, Title VII funds provided support for the program; for the remainder of the study only State bilingual funds were used in support of this effort. The primary goal of the bilingual program is to prepare Limited English Proficient (LEP) students to function in an all-English instructional program. At kindergarten, the children are heterogeneously grouped. Bilingual teachers provide instruction to both English proficient and LEP students. The former receive all of their instruction in English. ESL is provided for LEP students. Spanish is used to assist the LEP students in learning concepts, and reading readiness in Spanish is also provided for these children. At first grade, English proficient students are assigned to the regular all-English school program; LEP students are assigned to the bilingual classroom where they receive reading instruction in Spanish and bilingual support in other subject matter areas. ESL is also provided. These children remain in the bilingual program until such time that they have gained reading skill in Spanish and sufficient oral English skills to function in an all-English curriculum. Children may remain in this program through fifth grade, however, approximately one-half of the students in the study who were assigned to the bilingual program at first grade were mainstreamed by third grade. None of the target students who were fourth graders at the close of the study were enrolled in bilingual classes. The principal of the school administers the bilingual program. During the two years that the district had Title VII funds, a coordinator was employed to assist with testing and placement of students, planning and carrying out staff development activities, matters of curriculum, and general supervision of the bilingual program staff.

School Sites

The target students from this site were drawn initially from the district's primary school which houses kindergarten through third grade and serves some 426 students at those grade levels. The classes are classrooms opening into a central courtyard. Adequate playground space surrounds the building but because of the sandy composition of the soil, grassy areas are limited. At times the weather conditions are such that the sand prevents the children from playing outdoors.

On completion of third grade, the students were transferred to the district's elementary school which serves some 448 students in grades four through six. The classes are housed in the main building and in three or so temporary buildings adjacent to the main building. The main building is similar in structure, style, and age to the primary school building.

Administrative/Supervisory Personnel. The principal serves as the instructional leader at the building level. The principal at the primary school has been in that position for a number of years and provides strong leadership in the reading program at that school. The principal at the elementary school has also been in the district for several years and was assigned to the principalship of the elementary school for the last two years of the study.

At the beginning of the second year (of four) of the study, a new superintendent was hired, due to illness of the former superintendent. At that point, two major changes were made in the structure and nature of the reading program. These are described below in subsequent sections.

Teaching Staff. Approximately 103 teachers are employed in the school district. At the primary and elementary schools, approximately one-fourth of the teachers are Hispanic and fluent in Spanish. There is a moderate turnover of teachers from year to year. Some of the teachers live in the community, others live in nearby towns, and still others commute from El Paso. Depending on funding, a small number of aides are employed; these are generally drawn from the community.

Classrooms. All classrooms were self-contained. During the first two years of the study, children eligible for Title I (Chapter I) or Title I Migrant programs received additional reading and oral language instruction during pull-out classes. At the beginning of the third year of the study, these programs were restructured. A team of two teachers for each grade level was hired for these programs. The instruction was then carried out in the regular self-contained classroom with the regular teacher working with the team for one hour per day. A typical class arrangement was for the self-contained classroom teacher to teach her/his regular reading class, then at another period the reading team would come into the classroom. The regular teacher would take about 10 or 11 of her students who did not need help and work with them on enrichment, reference, etc.; the Title I teacher would take about eight eligible students and work with them while the Migrant teacher would take her four or five students for instruction. The Special Education LLD teacher would take students from that classroom eligible for her/his service also at that time. These teachers work in cooperation with the regular teacher, using the basic reading textbooks with supplements. This restructuring of these special

89

programs was one of the two major changes referred to above that was implemented by the new administration.

The classrooms at each grade level were supplied with essentially the same basic curriculum materials, which in general appeared to be adequate. The curriculum materials remained relatively stable throughout the course of the study with the exception of the following: a new basal reading series in English was adopted and implemented during the 1980-1981 school year, and a new edition of the previously-used Spanish basal reading series was purchased.

Class size ranged from 23 to 28 students. The training and experience of the teaching staff varied, however, several of the teachers were long-termed teachers in the district. Most of the teachers new to the district during the course of the study brought with them previous teaching experience. The teaching staff assigned to the bilingual classrooms over the course of the study numbered five teachers; all were Hispanics and fluent in Spanish. Two of these, one kindergarten and the first grade teacher, left to take positions in a neighboring school district at the end of the first year. The new teacher hired for the first grade bilingual classroom for the second year of the study had taught previously at the secondary level in another school district but did not have elementary school teaching experience. He was at that time, however, enrolled in elementary education courses at a local university and also received periodic inservice assistance from staff from the Texas Education Agency Region Service Center throughout the remainder of the study. Seven of the target students were with this teacher in both their first and third grades, and another seven were with him for their third grade year. Otherwise, the students had a different teacher each year, except for one target student from the English proficient group who was retained at first grade.

The Reading Program

During the first two years of the study, this district employed an "individualized" approach in reading instruction. Children were tested, their reading level determined, and instruction prescribed on the basis of individual need. Management was carried out through a system of student contracts. In each classroom, a variety of basal readers were available and utilized (e.g., Harcourt Brace; Ginn; Harper & Row; Scott Foresman; Holt, Rinehart & Winston) as were skill boxes containing a variety of supplementary materials. At the beginning of each week the student was givn a "contract" to follow in which assignments for each day were indicated; these were prepared by the teacher. The contracts contained codes that designated the materials to be used and the tasks to be undertaken. Each student then worked independently to complete each of her/his assignments. The teacher, and usually an aide, monitored the work and provided assistance. In addition, the teacher worked individually with the children and also frequently brought together small groups for direct instruction or evaluation of work. This approach was used at all grade levels in the primary school in both the English and Spanish components of the program.

Beginning in the third year of the study, with guidance from the new central administration, the "individualized" approach to instruction was discontinued and was replaced with the district's current program. In the present program, the district-adopted basal series serves as the basic materials for instruction. At each grade level, all children are instructed with the textbook designated for that particular grade level by the basal series (i.e., all third graders are instructed in the textbook that bears the designation 3^1 or 3^2). The teachers may supplement the instruction with other materials that are designated to be used above or below the particular grade level, but the student is expected to attempt work designated for her/his grade level and to ultimately achieve at that level during the course of the school year.

The reading classes are self-contained. The children receive direct instruction in small groups for approximately 20 minutes per day. The remaining portion of the reading period (usually one hour) is spent in independent work, usually based on the basal series workbook, ditto sheets, or chalkboard assignments. The homeroom teacher is responsible for the instruction in all other subject areas except P.E. and Music, which is usually taught by specialists. As described in a previous section above, children eligible for Title I, Title I Migrant, or Special Education LLD classes receive additional reading and language instruction.

The basal series currently in use in the English reading program is the Houghton-Miflin series. The Santillana Publishing Company materials serve as the basic materials for the Spanish reading program.

Students in the Spanish reading component of the program are evaluated periodically during the course of the school year. When they reach "transfer criteria" they are assigned to the all-English, mainstream instructional program.

Transfer Criteria:

- scores at the designated level on the oral language proficiency test
- scores at or above the 40th percentile on the California Achievement Test

Students scoring in the range of 23rd-39th percentile may be transferred to the mainstream program on positive recommendation from the bilingual classroom teacher.

Staff Development

The school district typically provides five days of inservice training during the school year. Four of these usually occur just prior to the opening of school in the fall; the other usually occurs at about mid-point during the year. Topics for the inservice days are based on identified needs. Consultants are sometimes brought in from the local regional service center, from neighboring school districts, or from the Bi-County Cooperative, which provides inservice training in special education. At other times, personnel from within the school district (including teachers) who have expertise in an identified topic conduct the inservice sessions. Typical of the topics included in the last two years are orientation to new materials or methods adopted by the district, classroom organization and management, reading instruction, and objectives related to the Texas Assessment of Basic Skills instrument.

Some Problems and Concerns

Problems and concerns expressed to the research staff during the course of the study center on three topics. The inadequacy and accuracy of the available language proficiency tests gives rise to concern about placement and termination of special language-assistance services for bilingual students. The use of standardized achievement tests in English with this population, in the absence of availability of such tests in Spanish, is thought by some to underestimate the academic achievement of students in these programs. Finally, while the district has been successful in recruiting teachers, there is a desire to have the teacher training institutions in the area place teacher trainees from their community (i.e., preservice teachers planning to return to the community to teach) in their schools for their intern period. This would present some obvious advantages both to the preservice teacher and to the district as well.

SEDL BILINGUAL READING STUDY

Summary Description of Project Sites

Central Texas - Site 3

District Setting

The district is located in central Texas, 51 miles northeast of San Antonio. The general characteristics of the school district, as of 1982, are listed below:

Size: Schools - 6 (3 elementary, one each serving pre-K-1, 2-3, 4-5; 1 intermediate, grade 6; 1 junior high, 7-8; 1 high school, 9-12).

Students - 4,615

Teachers - 270

SES (community): low to lower middle income households

Ethnic Composition:

Hispanic	64.9%
Anglo	30.9%
Black	04.1%

Distribution of Hispanic Population: Hispanics are enrolled in all six schools and at each grade; they reflect the district's ethnic composition.

Level of Support - ADA: 4,574

Per Pupil Expenditure: $1,842

Enrollment Trends: Relatively stable; slight increase in recent years.

The school district is situated in a small, semi-urban community in south central Texas. A few small factories and a state-supported university serve as the economic base for the community. Approximately 41% of the community's population is Mexican American, and more than 59% of these families earn an annual income below the national poverty level. The Hispanic population in the local district is concentrated on the south side of town; approximately one-half of the Mexican Americans live in this densely populated medium to low income area.

The school district serves some 4,615 students in pre-kindergarten through grade 12. The school population, of which approximately 65% are Hispanics, has been relatively stable, with only a slight increase over the past few years. The student population is distributed among six campuses: three elementary schools (pre-K-5); one intermediate school (grade 6); one junior high school (7-8); one high school (9-12). Each elementary school houses two grade levels only: Pre-K-1, 2-3, 4-5.

The Bilingual Program

Bilingual education was begun in the district in the 1968-1969 school year through a Title VII project which was granted to the local university. It was initiated in grades K-3 as a demonstration project for a three-year period. A second Title VII project was funded to extend this project for a two-year period. At the end of this five-year project, it was decided that Title VII funds for this project would no longer be sought. A local language enrichment program was modified and implemented at that time which included four components: Spanish for Spanish Speakers, Spanish-as-a-Second-Language, English-as-a-Second-Language, and English Language Enrichment. This program has been in grades K-5 since that time.

In 1973, bilingual education for LEP students was mandated by State law for students in kindergarten through third grade, and State monies were made available for this purpose. Beginning in the 1980-1981 school year, these services were extended to include students through grade six, in keeping with subsequent state legislation. ESL instruction for LEP students in grades 7-12 was also begun at this time.

During the school year of 1981-1982, services being provided to LEP students were not only supported with State bilingual funds, but also with three federal grants. These grants included: (1) a basic grant under Title VII for four-year-old pre-kindergarten students; (2) a basic grant funded under Title VII for students in grades 2-5 and (3) a one-year grant under Title IV to assist in the overall implementation and coordination of the district's Lau compliance plan.

The bilingual program in pre-kindergarten through grade three provides instruction to all children in their home language in the content areas of language arts, math, science and social studies as well as language development in their second language. Thus, Spanish dominant LEP students receive instruction in Spanish in the above content areas with 30 minutes per day devoted to ESL instruction. English-dominant

LEP students are instructed in the above content areas in English and in addition receive 30 minutes per day in each of the areas of ESL and Spanish for Spanish Speakers. All English proficient students, including monolingual English speakers, receive 30 minutes per day of Spanish-as-a-Second-Language in addition to instruction in the content areas in English.

The objectives of the program are to (1) assist LEP students in learning academic concepts through their dominant language while obtaining the necessary proficiency in English to make the transition to an all-English instructional program and (2) promote cultural acceptance and diversity by having all the children exposed to both languages and cultures. While the children are grouped by language category for a portion of their instruction, they are heterogeneously grouped for other activities (e.g., P.E., art, music, playground and lunch periods), giving them exposure to a wide diversity of cultural and language experiences through contact with their peers.

Only the ESL component of the program has been implemented for LEP students at the fourth and fifth grades. The instructional program for these students is the same as for non-LEP students, except for 30 minutes of instruction per day in ESL three days per week.

The program is administered through the district's Bilingual Programs Office which is staffed by a bilingual programs director, secretarial personnel, and depending on funding, a full-time coordinator who assists at those campuses where there is the greatest need. The present bilingual programs director has held this position for a number of years. She is a well-informed, dedicated leader with classroom teaching experience and administrative training. She strongly supports the concept of bilingual-bicultural education and the involvement of parents in the education of their students. She attends and participates in many meetings of the central administration personnel and periodically reports directly to the school board on matters involving the education of language-minority children in the district. Much of her time is spent in dealing with administrative details related to state and federal funding agency requirements and in working with the school staff in matters of staff development and recruitment of teachers.

The Bilingual Programs Office maintains a cooperative arrangement with the local state university in the city in the training of bilingual programs personnel at the inservice and preservice level, and in the training of student teachers. Bilingual teachers are often recruited from this program, as well as from other teacher training programs in the region.

School Sites

The target students at this site were drawn from the kindergarten population at the pre-   school. On completion of first grade, these students were transferred to the grades 2-3 elementary school.

The Pre-K-1 School

The open classroom structure at this campus involves large, open spaces in which "units" of approximately 100 children are served by eight teachers and five aides who work together to provide individual and group instruction to both LEP and non-LEP students. The target students for this study were housed in three such units.

The total population of the school numbers some 743 students of which some 67% are Hispanic. Of these, approximately 53% of the K-1 students are classified as Limited English Proficient.

The classes are housed in a modern, one-story, brick building that has several wings, each opening out into courtyards or play space. Some temporary structures connected to the main building house the four-year-old program and some of the support services. The instructional areas were well-lighted and comfortable; instructional materials were highly visible in work areas and on bulletin boards and wall spaces. The furniture in these areas consisted primarily of small tables and chairs which allowed for the flexibility of grouping needed to accommodate team teaching. Each wing (or unit) contained a teachers' work area which also housed a wide variety of well-organized and catalogued teaching materials (both commercial and teacher-made).

Administrative/Supervisory Staff. The principal serves as the instructional leader at the school level. The principal at this school has held this position for some time and has been largely responsible for implementing and developing the present open space, team teaching concept in the school. Within each "unit" of students, team leaders are designated for each curriculum area. Each team leader is responsible for providing leadership in her curriculum area within the team and for working with her colleagues with a similar designation in the other units. Turnover among this staff was minimal during the course of the study.

Teaching Staff. Some 40 people comprise the instructional staff at this school. Of these, about one-third are Hispanic. The teachers assigned to the Spanish component of the program within each unit hold the State of Texas Bilingual Endorsement certification. Turnover of teachers was minimal throughout the course of the study.

Instructional Areas. The teachers worked with small groups of children, with individuals, and at times with large groups in the open areas. A system of scheduling and movement of students was operationalized so that there was a minimum of confusion and time involved in changing instructional periods. The teachers planned together daily and regrouped students frequently on the basis of instructional need. The Title I and Title I Migrant teachers were assigned to each unit and served eligible children within the framework of the team-teaching plan.

The Grade 2-3 Elementary School

The total population of the school numbers some 653 students of which some 60% are Hispanics. Of these, approximately 42% are classified as Limited English Proficient.

The classes are housed in a modern, one-story brick building of a similar external structure of that of the Pre-K-1 building, however, walls are retained that divide the building into classroom units. The classrooms were equipped with either individual desks or tables and chairs to accommodate small groups of students. They were well-lighted and comfortable; instructional materials and brightly-decorated bulletin boards were readily visible.

The organizational structure of the school differs considerably from that of the Pre-K-1 school. English proficient children are heterogeneously assigned to homerooms of approximately 30 children. For Reading instruction, however, children in the English reading program are grouped by reading achievement and are scheduled for a period of reading instruction (approximately one hour daily) with a specified teacher for a particular reading level. Limited English Proficient students are assigned to a bilingual classroom and scheduled for a period of reading instruction in Spanish (usually one hour) with a Spanish reading teacher.

Administrative/Supervisory Staff. In addition to the principal, who serves as the instructional leader in the building, the school had the full-time services of a coordinator for the bilingual program. The role of the coordinator included assessment and placement of students, record keeping related to school district and funding source requirements, and working directly with school personnel in matters of curriculum and in planning and carrying out staff development activities.

Teaching Staff. An instructional staff of some 38 people were employed at the school. Approximately one-third were Hispanic. the teachers assigned to the Spanish component of the program hold the State of Texas Bilingual Endorsement certification. Title I and Title I Migrant teachers served eligible students in pull-out classes.

The Reading Program

The reading program in the school district incorporates both a management system and a basal reading series in the English component and a basal series in Spanish components. Monolingual English-speaking, English proficient bilingual, and English dominant Limited English Proficient students receive reading instruction in English only. The development of prereading skills is begun in kindergarten (as well as for those children enrolled in the Pre-K program); for those children who make sufficient progress, formal reading instruction also begins at kindergarten. Montessori and Wisconsin Design materials provide the framework of skills to be taught. The early books of the district-adopted basal series (Houghton Mifflin), along with a variety of supplementary materials, provide the text for instruction. The basal series and the Wisconsin Design materials are the primary materials used in subsequent grades.

Spanish dominant children of limited English proficiency begin their initial reading instruction in Spanish. Montessori materials and the Spanish Reading Keys (Economy Series) are the primary materials used in Pre-K and kindergarten. A variety of basal readers are available from first grade on: Laidlaw, BoLAR, Santillana, and Economy Series. Formal instruction in reading may be begun for some of these children in kindergarten; most are in first grade before other than prereading skills are introduced.

As indicated above, kindergarten and first grade classes are provided in an open-space, team teaching structure. Instruction occurs in small groups, and children are regrouped frequently to accommodate special needs. At second grade reading instruction occurs in class sizes of approximately 30 students formed on the basis of achievement. Within each class, further grouping occurs on the basis of need.

Transfer Criteria

Children who begin their initial reading instruction remain in the bilingual program until they meet specified criteria which are reflected below in the district's guidelines for transition from Spanish to English reading instruction:

1. The child should have a minimum of one year of reading instruction in his native language.

2. The child should have a minimum of one year of language development in the second language (ESL).

3. The child should be able to master 80% of the SOLD (System for Oral Language Development), Level 6, Placement test.

4. The child should be able to answer 80% of the following comprehension questions after the teacher reads "The Painted House" (pp. 104-110) from People and Places (Harcourt Brace Jovanovich):

   - Name the characters in the story.
   - Why was the house being painted?
   - How do you know mother wasn't angry at father?
   - Describe some of the paintings on the house.
   - How did the story end?

5. Score of 3 or above on the Language Assessment Scales (English).

Staff Development

In addition to the district's "regular" inservice program for all teachers (approximately five days during the year) the bilingual program teachers typically are provided a variety of other inservice activities. These include workshops, in-classroom assistance by program staff and outside consultants, participation in regional conferences, and curriculum writing. In any given year the bilingual program teachers may accumulate as much as the equivalent of 15 days of inservice training. Examples of the topics of workshops provided by outside consultants are: Strategies for ESL, How to Administer and Score the Language Assessment Scales (LAS), Spanish-as-a-Second Language, Bilingual Education and English-as-a-Second Language, Baile Folklorico Dance Workshop, Selecting an Appropriate Spanish Basal Reader, Discussing Spanish Language Arts - How Can We Improve?, ESL - ›bal Approach, English-as-a-Second Language Methods and Techni , Language Acquisition: A Process Overview, and Grade and Cour.esy Lessons.

Some Problems and Concerns

The problems and concerns expressed to the research staff by district personnel during the course of the study included the following:

1. Identification, diagnosis, and program placement of Limited English Proficient (LEP) students. A number of problems are related to this concern. First, most LEP students in this district are English dominant, but they are not a homogeneous group. Some come from homes where the parents themselves are Limited English Proficient but who speak primarily English to their children; others come from homes where the parents are Limited English Proficient but who speak primarily Spanish in the home. Assessment and diagnostic tools simply are not available that provide the kind of extensive and accurate information that is needed on which to make instructional decisions relative to these children. Nor is sufficient numbers of adequately trained assessment personnel available to administer those instruments that do exist. Consequently, traditional reme al programs are provided to these children who, in the opinion of some, do not address the real needs of these children.

2. Teacher training. Preservice training of teachers for the bilingual program does not equip them to carry out appropriate and accurate assessment, to interpret results of assessment, nor to diagnose needs and plan appropriate instructional treatment for language minority children. The training of teachers for the regular, mainstream programs does not equip them to work effectively with language minority children.

3. Performance on standardized tests of academic achievement. Language minority children in this district are generally relatively successful on these tests at the end of first grade and, on the basis of this, are often transferred to the regular, mainstream program. However, by the end of third grade, the performance on these tests is low. In the

opinion of some, the heaver language demands, and the kind of language demands, placed on children as they proceed through school, as opposed to that required for kindergarten and first grade, requires special language assistance programs beyond first grade for many of these children.

4. The junior high and high school dropout rate of children who were English dominant LEP children on entry into school is higher in this district that that of the general school population.

SEDL BILINGUAL READING STUDY

Summary Description of Project Sites

Northern Mexico - Site 4

District Setting

The district is located in an isolated, middle-sized Mexican city south of El Paso, Texas, some 200 miles from the US-Mexican border. The general characteristics of the school district, as of 1982, are listed below:

Size: Schools - 8 (Federal)
25 (State)

Students - 19,800 (approximate)

Teachers - 396 (approximate)

SES (families served by the schools in the study):
Federal - low to lower middle
State - lower middle to middle

Ethnic Composition: Mexican (100%)

Distribution of Hispanic Population: Hispanics make up the entire school population.

Level of Support - ADA: 18,810

Per Pupil Expenditure: 17,000 pesos ($100.00)

Enrollment Trends: Federal - increasing yearly
State - decreasing yearly

Brief Overview of Mexican Educational System

As is the case in the United States, the Republic of Mexico is divided into geographical units referred to as "states." Each state elects a governing body and a governor; similarly each town and city has its set of elected officials, including a mayor.

The federal government is housed in and operates out of the federal district located in Mexico City. It is governed by a President who is elected and serves one six-year term. Matters of education throughout Mexico come under the auspices of the Ministry of Education, a department of the federal government housed in Mexico City. This department is headed by a Minister of Education who is appointed by the incoming President. The Minister of Education usually serves the concomitant six-year period with the President who appointed him. This is a political appointment, and as is often the case with political appointments in most countries, the preparation and background of the Minister of Education in matters of education vary from one administration to another.

Federal monies spent on education are administered through the Ministry of Education. Curriculum, for both private and public education, is set in the Ministry. Textbooks are adopted centrally and are published by and distributed from that source. The Ministry of Education sets rules and regulations; schools, both private and public, are monitored and supervised rather rigidly to ensure adherence to those rules and regulations.

At least four separate and distinct school systems operate in Mexico. Two provide "free" education (Federal and State systems); two are tuition-based (parochial/ethnic and private systems). In all cases, the federally-decreed Spanish language curriculum is required. Foreign language instruction is typically offered in the parochial/ethnic and private schools, and some of these are fully bilingual schools where the curriculum is taught in Spanish and one other language.

Federal system. This consists of schools, widely spread throughout the country, that are supported completely by funds from the federal government. Teachers for these schools are often recruited and trained in normal schools (teacher training institutions) supported by the federal government and are subsequently hired by the federal government and sent out to schools wherever they are needed. The Ministry of Education maintains close control over and supervision of these schools.

State system. This consists of schools established at the state level and supported primarily by state funding with only a percentage of the funding coming from federal sources. The State has full responsibility for the supervision of these schools, but they are, nonetheless, subject to the rules and regulations set forth by the Ministry of Education in matters relating to, for example, curriculum, length of the school day, and the training of teachers. The states do, however, have the opportunity to go beyond the minimal standards set forth by the federal government and to, in certain ways, tailor education to their local needs.

Parochial/Ethnic schools. Some of these schools are owned and operated by nuns and priests. They are not, however, allowed to include religious teachings in their curriculum. Some are Jewish schools, attended by children of Jewish families, but no religious training is permitted. Families pay tuition for their children to attend, but these schools, nonetheless, are subject to all the rules and regulations of the Ministry of Education, including required use of the federally-approved Spanish language curriculum and textbooks.

Private schools (of various kinds). These are owned and operated by individuals or corporations, and tuition is usually substantial. Examples of these include:

1. Montessori schools - offer early childhood education as well as schooling through the elementary grades.

2. Cooperative schools - offer elementary and secondary education. They are owned by the parents of the students. The children who attend pay a subscription fee plus tuition on a regular basis.

3. International/Bilingual schools - children in these schools receive their education in two languages. For example, the American School provides bilingual schooling in English and Spanish. Others provide schooling in French and Spanish or in German and Spanish.

4. Institutes - these "schools" typically provide special classes in English as a foreign language and Spanish classes for foreigners. Other kinds of classes may be offered as well.

5. Technical schools - these are oriented toward job-related skills. Tuition is usually required, although particular schools may receive government support of one sort or another.

In Mexico class lines, both social and economic, are sharply drawn. Federal and State schools are usually attended by children of lower and lower-middle SES families. Middle and upper-middle SES families pay tuition for their children either in parochial/ethnic or private schools. Upper SES families generally send their children to private schools. Some choose Spanish language schools, but often as not, their children are educated bilingually in French, German, or English. Others send their children to private residential schools in the United States, Switzerland, or France.

The facilities and quality of instruction vary considerably among the various types of schools. However, regardless of the type of school, differences in educational practices between those in the United States and those in Mexico, at any given time, are apparent. For example, class size is usually much larger in Mexico; the curriculum and teaching practices are more uniform than they are in the United States, and the facilities, extracurricular activities and services are less elaborate and extensive than they are in many parts of the United States.

School Sites

The target students at this site were drawn from the first grade classes at two schools. One school (School A) is a state-supported school and serves approximately 700 students in grades 1-6; the other (School B) is a federally-funded school which serves some 712 students in grades 1-6. There were two first grade classes at each school; 15 children were selected from each first grade class (60 students).

Both schools are located within the city boundary. Both serve primarily monolingual Spanish-speaking children. School A serves children from lower middle to middle income families; children in School B come from low to lower middle income families. The children attended school for four and one-half hours per day, from early September through late June with several holidays and rather long breaks at Christmas and Easter. The school day normally is approximately four and one-half hours long (from 9:00 AM to 1:30 PM or 8:30 AM to 1:00 PM). The children go home for the mid-day meal and do not return. However, because of limited facilities and the large number of children to be served by the schools at this time in this site, two shifts of students were being served (8:00 AM to 12:30 AM and 12:30 PM to 5:00 PM). Each shift was taught by a different teacher. The classes were housed in low, modern, brick structures in good repair. They were well-lighted, and usually comfortable except during extreme cold spells when heating, not normally needed, was desirable. School A classrooms were equipped with rows of student desks, each shared by two or more children. School B classrooms were equipped with small tables and chairs and were arranged in clusters to accommodate individual and small-group instruction.

The classes were large (approximately 50 students per class). The classes were self-contained, and all instruction was provided by a teacher who had completed normal-school training.

All of the classes used the federally-adopted textbooks, a set supplied for each child, in which all subject matter to be taught is integrated into one set of books (i.e., math, science, social studies and the language arts are interwoven into the same set of textbooks).

The director (principal) of the school was responsible for the management and administration of the school. Supervision and monitoring of the instruction was carried out on a regular basis by outside personnel under the auspices of the state and/or federal government.

The Reading Program

Since teachers must follow the national guidelines and use the government-approved textbooks, the content and methodology of the reading instruction was quite similar in both of the schools. However, the organization of the students and the delivery of the instruction differed between the two schools. In School A (the state-funded school), most of the instruction was presented to the full group and was characterized by much direction on the part of the teacher and choral response on the part of the students. In School B (the federally-supported school) the teachers were involved in experimenting with a delivery system known as "The Workshop Way" in which direct instruction occurred both in small groups and on an individual basis. The children were guided through a series of activities daily in which they proceeded at their own pace and received help as needed either from the teacher or from a peer. While the work was individualized to a certain extent, each child was expected to complete her/his assigned tasks daily and was responsible for soliciting aid when needed.

The approach to reading instruction in the schools is referred to as the "Global Method of Structural Analysis," which is a four-stage approach to developing reading and writing skills. The four stages are: (1) visualization of utterances, (2) analysis of words themselves, (3) breakdown of words into syllables, and (4) affirmation of previous instruction in reading and writing whereby comprehension is induced using all elements which make up an utterance.

In the first stage, after having engaged the students in general conversation or conversation directed toward the content of the readings, the teacher chooses several of the utterances to write on the board. The teacher then reads these aloud and directs the students to read them aloud along with the teacher. Associating written and spoken language, the teacher asks the students to identify the written utterances by having them tell "what they say." Also, the teacher randomly selects from the utterances written on the board, reads the utterance, and asks the students to indicate which utterance was read. The students then copy the target utterance into their notebooks. The purpose of this is to develop the students' knowledge of the relationship between speech, reading, and writing.

The second stage repeats the activities of the first. The teacher follows these activities by focusing instruction on the individual words which comprise the utterances that are on the board. The teacher reads the words separately and individually, and then the students repeat them along with the teacher. Then, pointing randomly to individual words, the teacher asks "what they say." Subsequently the students copy several words in their notebooks and in some way indicate their meaning. They then copy each utterance, one by one, into their notebooks and indicate its meaning.

In the third stage, the students are instructed in the analysis of words by the vowels they contain. Since vowels are the syllable nuclei, knowledge of them is necessary in order to analyze the syllable. The sequence of presentation of both vowels and consonants is based on their decreasing frequency of use in Mexican Spanish. Along with frequency, the degree of difficulty of the sound-symbol relationship is considered. Moreover, letters with similar physical features are separated in the sequence so as to avoid the problem of visual discrimination.

Instruction on the vowels is a prerequisite to achieving the goal of the third stage which is to instruct the students in syllabic structure. In the third stage, the teacher repeats the activities of the first two stages and proceeds by isolating one type of syllable for in-depth study. The teacher asks the students to identify the target syllable in the words which have been written on the board. After writing several of the words in their notebooks, the students are asked to underline or highlight the specific syllable. Continuing, they are shown the various consonants or combinations of consonants that may appear in the syllable, as well as other vowels which may occur, by illustrating them in words in their readings or in words already on the board. Next, the teacher has the students form words by combining different syllable types. Finally, the students orally express utterances which contain some of the words which they had formed earlier and then copy them, along with the words on the board, into their notebooks.

The fourth and final stage is an affirmation of the first three. Again the teacher engages the students in conversation in order to elicit utterances for study. These are written on the bord and then read aloud by the entire class. At this point, comments are made regrding the thematic content of either the utteranes or of the readings. The students write in their notebooks some of the utterances which relate to the theme of the reading or the general theme of the chapter in which the reading appears. The readings are read aloud by the teacher and the students so that (1) the students become familiar with the content and (2) their knowledge about the relationship between speech and print is reinforced.

Staff Development

Staff development activities occurred at various times during the course of the study. In School A, the first and second grade teachers had six workshops each year on teaching methodologies, while the third through sixth grade teachers had two seminars yearly on this topic. At School B, the first and second grade teachers had four workshops per year on teaching methodologies, while the third through sixth grade teachers were provided two seminars on this topic. In addition, four workshops on the teaching of reading were provided yearly for the first and second grade teachers in School B.

Some Problems and Concerns

Concerns expressed by members of the local research staff during the course of the study were related to the following:

1. The amount and kind of preservice training required for public school teachers.

2. Large class size.

3. Shortening of the school day to accommodate two shifts of children and limitations on the use of the facilities and materials due to the necessity to serve two shifts (e.g., teaching materials could not be displayed and left in the room from one shift to another; children did not have access to the textbook materials except during their actual instructional time).

4. Irregular attendance on the part of some children and the instability of a given student population from one year to another.

5. Lack of variety in and quantity of the teaching materials.

SEDL BILINGUAL READING STUDY

Summary Description of Project Sites

East Texas Area -- Site 5

District Setting

The district is located in a large urban area in northeast Texas. The general characteristics of the school district, as of 1982, are listed below:

Size: Schools - 99 (69 elementary; 18 middle; 12 high schools)

Students - 65,125

Teachers - 3,069

SES (community): low income households (less than $15,000) - 44%; middle or upper income households - 56%

Ethnic Composition:

City		District	
Black	22.63%	Black	36.3%
Hispanic	12.64%	Hispanic	19.7%
Anglo	63.49%	Anglo	42.3%
Other	1.24%	Other	1.8%

Distribution of Hispanic Population: Concentrated in 18 elementary, 3 middle, and 1 high school in high density, low SES, inner-city areas.

Level of Support - ADA: 57,941

Per Pupil Expenditure: $2,724

Enrollment Trends: Declining enrollment of overall student population (15% since 1973; decline less severe in last 3 years).

Increase in Hispanic population (from 10.1% in 1971 to 19.7% in 1982).

The school district serves grade levels kindergarten through 12. The district has been experiencing a decreasing enrollment over the past few years. In 1971, the total enrollment in the district was 79,494 and had dropped, by Fall 1982, to 65,125. During this same period the percentage of Hispanic students has steadily increased from 10.1% in October 1971, to 19.7% by October 1982.

Currently (1982-1983), there are 69 elementary schools, 18 middle schools and 12 high schools in the district. The Hispanic population tends to be concentrated in the 18 elementary schools, three middle schools, and one high school served by the district's bilingual program.

Bilingual Education Program

The district provides a special language assistance program in grades K-12 to serve limited English proficient (LEP) students. For kindergarten through grade 5, it is a full-time educational program designed to allow students to learn academic concepts in their home language while obtaining proficiency in the English language. The ultimate goal of the program is successful academic achievement by all students in an all-English curriculum program. The history and culture associated with both languages is emphasized to instill a sense of pride and identity in the students.

In the 1982-1983 school year, bilingual education was in operation in 26 schools in grades K-5. This included the 18 elementary schools with high concentrations of Hispanic students as well as eight additional schools where sizeable pockets of such students were located. The bilingual education classrooms include 23 kindergarten, 29 first, 19 second, 19 third, 9 fourth, and 9 combination classrooms in third, fourth, and fifth grades.

The secondary program (grades 6-12) for LEP students is an English-as-a-Second-Language (ESL) program. Bil ngual aides assist the content area teacher at each of three secondary schools: one high school and two middle schools.

The program is administered through the district's Bilingual Programs Office which is staffed by a bilingual programs director, one coordinator, one instructional assistant, two staff development specialists, two bilingual reading specialists, one language laboratory specialist, and secretarial personnel. The "specialists" staff is funded through Title VII monies. The bilingual programs director, formerly a classroom teacher in the district, has held the position for several years. She is a well-informed, dedicated leader who believes in use of the home language of bilingual children as the bridge to mastery of English and ultimately to success in the regular mainstream classes. She attends and participates in many meetings of the central administration personnel and periodically reports directly to the school board on matters involving the education of language-minority children in the district. An active and vocal segment of the Hispanic community provide input into matters related to the improvement of schooling for minority youth. The program director is often placed in the position of trying to work out ways of implementing the wishes of the community within the framework of existing administrative policies and fiscal contraints. In addition, she spends much of her time in administrative details related to state and federal funding agency requirements and supervising the staff of the Bilingual Programs Office. The supervisory staff works directly with school personnel in planning and carrying out staff development activities, matters of curriculum, testing and interpreting test data, and in diagnosis and placement of students.

The Bilingual Programs Office maintains a cooperative arrangement with several of the local universities in the city in the training of bilingual programs personnel, both at the inservice and preservice level, and in the training of student teachers. Bilingual teachers are often recruited from these programs, as well as from other teacher training programs in the region.

School Sites

Seven schools (A-G, below) were the home schools for the sample population for the students' kindergarten and first grade years. As a part of the district's desegregation plan, these students were assigned to three other schools (H-J) for their second grade year. The general characteristics of the 10 sample schools for the 1982-1983 school year are shown below:

			Ethnicity - %			
School	Size	SES	Other	Black	Hispanic	Anglo
A	474	72	1.2	4.3	84.8	5.7
B	143	54	0.0	0.0	96.5	3.5
C	371	40	.3	17.0	31.8	50.9
D	[illegible]45	[illegible]2	[illegible]	8.2	86.1	5.5
E	528	65	.2	10.8	68.4	20.6
F	426	66	.5	7.3	73.2	19.0
G	358	81	.8	22.9	65.4	10.9
H	88[illegible]	73	.3	50.3	42.6	6.8
I	697	34	1.3	48.9	7.3	42.5
J	384	75	0.0	35.7	46.9	17.4

The principal is the instructional leader at the school level, however, the role and extent of the principal's direct involvement with the bilingual program within the building appears to vary from school to school. In three of the target schools with the highest concentration of Hispanic students (served by two principals), the principals appeared to be strongly supportive of their bilingual programs and were aware of the test scores and progress of the students. The research staff did not observe any outright opposition to the program by any of the school administrative staff in the target schools.

The ethnic composition of the schools varied, from 96.5% to 65.4% Hispanic in the home schools from which the sample population was drawn, with the exception of 31.8% Hispanic in the school selected from which to draw the monolingual English-speaking sample.

Schools A-F, with the exception of School C (containing the monolingual English-speaking sample), are located in the north end of the city, in close proximity to each other. The Hispanic population is concentrated in that area of the city. School G is located in the downtown area of the city, bordered on three sides by businesses and industry with one side of the school bordering on a low income residential area. This cluster of schools housed grades K-1 and 3-5. The children in the sample attended their K and 1st grade years in these schools.

As part of the district's desegregation plan all students at second grade are bussed to designated second grade schools where there is an effort made to include a racial mix of Black, Hispanic, and Anglo students in each of the schools. Our target students were assigned to three such schools (Schools H-J).

Schools A and B are adjacent to each other and are served by the same administrative staff. School B houses one of the district's preschool bilingual programs in addition to kindergarten classes. In both schools, all classes are self-contained. Some are designated as bilingual classes to serve limited English-speaking Hispanic students; these are staffed by bilingual teachers who provide instruction that follows the district's bilingual curriculum.

School C is located in the near western part of the city and serves a predominantly middle SES population. Only about one-third of the students are from Hispanic background. One classroom at each grade level (K-1) is designated as the bilingual classroom. Hispanic children of limited English-speaking ability are assigned to those classes. Each of these is staffed by a bilingual teacher who provides instruction according to the district's bilingual curriculum. The monolingual English-speaking sample for the study was drawn from the three English-medium kindergarten classrooms in that school.

School D is located in close proximity to schools A, B, E, and F. This school houses the district's four demonstration classrooms for the bilingual program. Teachers from other schools, as well as visitors to the district, often observe these classes. These classes also serve as a laboratory for evaluating materials and procedures. The classes in the building are self-contained. Hispanic children of limited English-speaking ability are assigned to one of the several bilingual classrooms in the school. The principal commented to the research staff on occasion during the course of the study that the demonstration program generates considerable pressure for the school staff. The teachers in the study, however, were most cooperative, confident in their work and showed a high degree of interest in the research.

Schools E and F are located in the same general area in the far north of the city and serve larger populations of Black and Anglo students than do schools A, B, and D. Certain of the classes in those schools were designated as bilingual classrooms within the school.

Schools H-J (Second grade schools) are located on the fringe of the attendance areas of schools A-G. In these schools, the children were assigned to homeroom classes. Limited English-speaking students were assigned to designated bilingual classrooms. The children were grouped by ability for reading instruction. Limited English-speaking students received instruction in reading in Spanish, and were grouped by ability for the Spanish reading classes (i.e., during the reading period one teacher taught children of a particular level; another taught children of a different reading level). In the bilingual homeroom classes, Spanish was used in the instruction in other curriculum areas. English- as-a-Second Language (ESL) was part of that curriculum. As children reached "transfer criteria," they were reassigned to transition reading classes, taught by one or more of the bilingual teachers. On successful completion of the transition curriculum, the children were then assigned to the regular mainstream reading program. All children were returned to their home school for the remainder of their elementary schooling, and proper placement, in either the regular program or bilingual program, was determined.

All bilingual classrooms in the study were stocked with essentially the same curriculum materials, which in general appeared to be adequate. The curriculum materials remained relatively stable throughout the course of the study. Program practices also remained relatively stable, however, there was one change of some significance. During the last year of the study, language laboratories were instituted in some of the target schools. Under the supervision of the Language Laboratory Specialist, paraprofessional aides provided ESL instruction to small groups of children (approximately 10 per group) whose English language skills were extremely limited. The Laboratory sessions, of approximately 30-45 minutes daily, supplemented the regular classroom instruction. In each of the sessions, the children rotated through different learning centers that focused on various skill areas (e.g., vocabulary, grammatical structures, following directions). Thus, the learning center approach allowed for a degree of individualized instruction.

All of the bilingual classrooms except two which were housed in portable temporary buildings, were housed in older, multi-story, traditional buildings, all of which were in good repair, clean, and uncluttered. The classrooms were well lighted and comfortable; instructional materials in both Spanish and English were highly visible in work areas and on bulletin boards and wall space.

Class size varied somewhat but was generally in the range of 20 to 30 students per class. In all cases, the teachers assigned to the bilingual classrooms held the State of Texas Bilingual Endorsement Certification. Most, but not all, were of Hispanic background. Throughout the course of the study there was a low turnover of the bilingual teaching staff in the target schools. Most of the teachers had had specific training for bilingual education, and the large majority of them had previous teaching experience in bilingual classrooms. The classrooms in general were well-managed and orderly.

The Reading Program

The district has adopted a "skills development" approach to reading instruction in both English and Spanish medium programs. Monolingual English-speaking and English proficient bilingual students receive reading instruction in English only. Limited English-speaking Hispanic students receive initial reading instruction in Spanish. In kindergarten and grade one, this instruction is delivered by classroom teachers, with supervision from instructional specialists, in self-contained classrooms. In the grade two schools, the students are assigned to homerooms, however, for reading instruction the students are grouped by ability and assigned to designated teachers for this instruction.

Basal reading series provide the foundation for both the English and Spanish reading programs. These series are structured so that the sequence of instruction is built into the readers and workbooks, which increase gradually in difficulty as the child progresses. The Macmillan reading series is the basic material for use in the English reading program. Materials currently in use in the Spanish component are The Spanish Reading Keys (Economy series). When the children who are enrolled initially in the Spanish reading component reach "transfer criteria," they are placed into a transition English reading program - Reading in Two Languages (Galloping, Level A and some continue in Lickety Split, Level B - Santillana Series). When the transition program is completed, the Instructional Specialist administers the English basal reading placement test (Macmillan) for determining reading level placement.

Transfer Criteria

In order for a child to be placed in the transitional reading program she/he must meet the following transfer criteria:

1. Passing score on the Mastery Reading test on Spanish 2^1 reading level (Mi Mundo) or a higher level.

2. Satisfactory English language proficiency test score administered by bilingual staff.

3. Positive teacher appraisal.

4. Positive Bilingual Instructional Specialist appraisal.

Staff Development

Staff development in this district is accomplished through a variety of means. In addition to the staff development activities for the general school staff, the Bilingual Programs Office provides additional staff development activities for the program personnel.

Two Bilingual Reading Specialists were assigned (1982-1983 school year) to a total of 18 elementary schools. These specialists assisted teachers in grouping students for reading, selecting and organizing reading materials, scheduling activities, determining transfer points, and implementing strategies to develop first and second language reading skills. In addition, the reading specialists assisted parents by providing activities and methods to help children at home with Spanish reading skills.

The district, with assistance from Title VII funds, maintains a Demonstration Staff Development Program to provide a controlled but realistic school setting for continued staff development. Four demonstration classrooms were identified, one at each grade K-3, in three schools. Two staff development specialists assisted the demonstration classroom teacher with classroom arrangement, use of current methods, scheduling, management and the use of materials. Bilingual classroom teachers that participate in this program attend and observe in demonstration class for one day during the school year. In addition, the staff development specialists work individually with them in the teacher's home classroom on specific needs.

The Bilingual Programs Office also conducts an extensive inservice training program each year in the form of workshops and sponsorship of attendance at conferences. During the 1982-1983 school year, teachers were involved in some 16 workshops and conferences that treated such topics as ESL and English language development, ESL materials, Spanish Reading Readiness, transition and high interest Spanish reading, oral language proficiency testing, and ESL in the content areas.

Some Problems and Concerns

Growing numbers of language-minority students to be served, who speak a variety of languages, along with the concomitant need for additional well-trained bilingual teachers, presents an interesting challenge for the school district. The mobility of the students within the district, as well as students moving in and out of the district also adds an additional challenge. The attrition rate of target students from this district was higher than in other districts in the study. Also the retention of target students at first grade was relatively high in one of the schools in the study.

Concerns expressed by district personnel during the course of the study were often related to (1) inadequacy of oral language proficiency tests and procedures; (2) quality of the instructional materials and current practices for this population, particularly in language development and reading; (3) the level of standardized achievement test performance of these students; and (4) need for adequate and appropriate evaluation criteria and procedures.

Final Report

TEACHING READING TO BILINGUAL CHILDREN STUDY

Volume 3
Measurement of Growth

Robert C. Calfee
Wesley A. Hoover
Betty J. Mace-Matluck

Document BRS-84-R. 1-III

Preston C. Kronkosky
Executive Director

Southwest Educational Development Laboratory
211 East Seventh Street
Austin, Texas
(512) 476-6861

November 1984

There were many individuals and institutions who contributed to this research effort. We wish to express our sincere gratitude to the parents, students, and school personnel who provided the necessary data from which this study is derived.

In addition, several other individuals made valuable contributions to the study, for which we are indebted: Robert C. Calfee, Sylvia C. Peña, and Blanca de Alvarez.

And finally, we wish to thank the local data collectors at the school sites, many of whom remained with the study throughout its duration: Ramiro Barrera, Beatrice Cantú, Irene Cavazos, Carolyn Cruz, María de Obregón, Gloria de Torres, Gigi Galván, Olga Hernández, Irene Méndez, Guadalupe Treviño, Rosalinda Villanpando, and Gloria Villarreal. Their patience, dedication, and hard work helped make this study a reality.

Betty J. Mace-Matluck
Wesley A. Hoover

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure	Page

PREFACE

In June 1978 the National Institute of Education (NIE) funded the Southwest Educational Development Laboratory (SEDL) to conduct a longitudinal study on the Teaching of Reading to Bilingual Children. Educators and policymakers alike have long recognized that the ability to read is essential for success in school, in work, and in life; yet many children from second-language backgrounds have trouble learning to read in schools today. The majority of these youngsters are from Spanish-language backgrounds and from low income families. Special programs designed to meet the needs of these children are provided in schools, but there is limited research evidence to guide the development, evaluation, and implementation of these programs. This study is intended to provide information that will result in greater insights into what constitutes a favorable learning environment for children from Spanish-language backgrounds, what instructional sequences and events promote successful and efficient learning of literacy skills, and what the language and literacy outcomes of current schooling practices are for a large sample of these youngsters.

The study was conducted during the years of 1978 through 1984. It is a comprehensive longitudinal investigation of the development of reading skills from kindergarten through fourth grade for a representative sample of more than 350 children from bilingual backgrounds, and for smaller samples of children who, on entry into school, were monolingual in English or Spanish. In this "natural variation" study, teaching and learning were carefully documented in field settings at the several sites.

The goals of the study were to (a) describe variations in both English and Spanish language ability of students living in bilingual communities, (b) document prevailing practices in reading instruction for bilingual students, and c) investigate the relations between the instructional program and student achievement for students with differing entry profiles.

Description of the Study

Surveys of the general and school populations reveal an increase in the number of students whose language resources are not an ideal match to the language of the school. An important question for educational practice and policy centers around the school's responsibilities in this situation. Bilingual programs, English-as-a-Second-Language classes, classroom aides, and "sink-or-swim" approaches can all be found in practice today. From limited evidence now available, none of these techniques has emerged as the one best system.

Hispanics make up the largest and fastest growing school-age population today. The demographics for some states show that over the next decade they may constitute as much as a third to a half of the population. In the state of Texas at present approximately one third of the school children are from Hispanic backgrounds (approaching one

million). They are found in virtually ever school district in the state. Many of the school districts in the southern portion of the state serve school populations of which 75% to 99% of the children are from Spanish-speaking backgrounds and, on entry into school, are often limited in their ability to speak English and to profit from instruction in that language. This population is not restricted to the border areas, however. Large urban centers in the state report as much as 20% of their school population from Hispanic backgrounds, with a concentration of some 80% to 90% in certain of their schools.

It is well documented that, in general, children from Spanish-speaking backgrounds, for whatever reason, often encounter difficulty in our nation's schools; they do more poorly on standardized tests than does the general school population, and their dropout rate is high. Bilingual education, in which students are given instruction partially through the home language until they have attained sufficient proficiency in English to benefit from English-medium instruction, has been the principal approach recommended by the Office for Civil Rights to ensure access to equal educational opportunity for these children. Although many individual programs have had considerable success in improving the academic performance of language-minority students, it has not been demonstrated that these programs generally are reducing inequality of educational opportunity on the large scale that was envisioned.

Growth in reading comes about for most youngsters through formal classroom instruction. Understanding the development of reading, and knowledge of the critical variables that determine success or failure, depends on a careful examination of the instructional program -- not just the label over the classroom door, but the program as actually implemented by the classroom teacher.

Educators have raised several issues about the most effective way to help bilingual children become proficient readers of English. These include (a) valid assessment of the student's ability in the languages of the home and of the school, (b) the optimal balance of formal instruction in both languages, (c) the most effective transfer from one language to the other, and (d) bilingual support within the classroom environment. A major thesis of the Teaching Reading to Bilingual Children study is that addressing these issues (and others) requires a comprehensive and ecologically-valid investigation of the linkage between the child's language and the language of instruction.

Design of the Study

To achieve the objectives of the study, considerable attention was given to the selection of schools, teachers and students, to the instruments for assessing language and reading achievement, and to the methods for evaluating the classroom instruction. Each of these topics is discussed briefly below.

Schools, Classes and Teachers

Twenty schools and 200 teachers from six school districts participated in the study. Included are variations in the nature of the reading program (a range from phonics-oriented to meaning-based), classroom organization (some self-contained, others team-taught), and grade structure (the range of grades in the individual school and the extent of cross-grading both vary). The schools differed in size, SES, urbanicity, locale, and makeup of the student body (from medium to high concentration of bilingual students).

Student Cohorts

The study was undertaken in four cohorts or "waves" of students. Three of the cohorts consisted entirely, or in large part, of bilingual students. The first cohort was small (N=40) and of limited generality; the second was somewhat larger (N=80) and covered a slightly broader array of contexts. The third cohort which was both larger (N=200) and broader in its generality, incorporated a number of procedural improvements based on previous experience in the study and included a monolingual English-speaking sample. The fourth cohort consisted of a relatively small sample (N=60) of monolingual Spanish-speaking students.

All of the bilingual sites were from the state of Texas, as were the monolingual English-speaking students. The monolingual Spanish-speaking students were from one site in Northern Mexico.

The original design of the study called for each student to be assessed and observed from entry to kindergarten through exit from third grade. By covering the full range of the primary years, we would be able to examine the transition from "learning to read" through "reading to learn." For students in programs where the initial stages of reading were in Spanish, we also considered it important to determine the transition to competence in English reading.

The original design was in fact implemented for the first two cohorts; some of the students were tracked from first through fourth grade, but most followed the intended design. Due to limited funding in the later stages of the study the last two cohorts could not be followed for the full four years that were originally intended. The bilingual and monolingual English samples from the Texas sites were observed from kindergarten through second grade, and the monolingual Spanish samples from the site in Northern Mexico were observed from first through third grade (the program did not provide a kindergarten).

The monolingual samples were incorporated in the design to aid in validating the instruments for student assessment. Both the English and Spanish cohorts are small and not selected to be fully representative of monolingual populations. Data from these samples will be presented in Volume 3, as part of the discussion on the adequacy of the instruments for measuring growth. The study was designed to study the course of reading in bilingual students, not as a basis for comparing these students with monolingual youngsters. Accordingly, comparisons

between the various samples will not be made in this report, nor do we reccommend that others attempt such comparisons.

Language Assessment

Several types of data were collected for each student on English and Spanish proficiency. Each year, early in the Fall and again in the Winter and Spring, teachers rated their students' language skills. Oral language proficiency tests were administered in the Fall of each year. Finally, audiotaped speech samples were obtained monthly on a rotating schedule in three settings: in the classroom, on the playground, and in the home.

Reading Assessment

Several instruments were used to measure reading achievement. Standardized test scores (mostly English) were collected yearly. More detailed information was obtained from a battery of individually-administered "performance based tests" in both English and Spanish. In kindergarten, the Stanford Foundation Skills Test was employed to measure the child's pre-reading skills. From the end of first grade on, the Interactive Reading Assessment System was administered during the Spring of each school year. This instrument provides independent measures of the student's skills in decoding, word meaning, fluency in oral reading, and comprehension. Finally, informal reading inventories were administered througho the school year.

Classroom Observations and Teacher Interviews

Project staff conducted monthly observations of the reading instruction in each classroom and interviewed the teachers quarterly about their instructional plans. The observation instrument documented staffing patterns, grouping and organization, time allocation, the language of instruction, the character of instruction, the materials and procedures used, and the response of the students. The interviews focused on the teacher's general instructional objectives, as well as the objectives for individual target students. Taken together, these two instruments yield a rich characterization of the classroom environment for the target students.

Student Entry Variables, Classroom Factors, and Reading Achievement

The primary goals of the analyses were to identify the general relationships that characterize variation in these factors and to look for underlying regularities that are associated with success and failure, both in the early stage of reading instruction and in the year-to-year variations.

Documents

This report is one of a series of eight documents contained in the Final Report submitted to the National Institute of Education. A com-

plete list of these documents is provided on the inside of the cover of this report.

The study was a collaborative effort among a number of individuals and institutions. All members of the research team contributed to the thinking, planning, and writing of this series of documents, however, the individual whose name appears first in the list of authors was responsible for preparing the particular document.

Betty J. Mace-Matluck
Wesley A. Hoover
Co-Principal Investigators

Austin, Texas
November 30, 1984

Introduction

This volume presents the methodology used in the analysis of the data from the study (a) to summarize patterns of growth in reading achievement, (b) to relate ancillary measures (language and prereading skills) to reading achievement, (c) to describe the instructional program during the primary grades, and (d) to examine the linkage between instruction and growth in achievement. The data from the first two cohorts will be used for illustration in this volume. The discussion will be organized according to the four tasks listed above.

Measurement of Growth

Measuring Growth in Reading

A major goal of the Teaching Reading to Bilingual Children study is the investigation of patterns of growth in reading achievement; growth is used as a generic term referring to changes in performance due to learning, development, or both. The discussion begins with a brief review of the concept of measurement of growth, and comments on the methods used in the study to measure growth in reading achievement. Next comes a presentation of the concept of a linear growth track as a means of summarizing the acquisition of reading skills. Then we will illustrate with the first two cohorts the results for average or aggregate measures of growth in the various areas of reading achievement that are tapped by the Interactive Reading Assessment System, following which examples of individual protocols will be displayed.

As will be apparent from the examples, the interpretation of the standard or "Y" intercept is subject to question. Accordingly, in this section of the volume we will present an alternative approach for describing the intercept of the linear function: the "X" intercept, which provides an estimate of the onset of instruction.

The examples also show that the individual growth tracks often contain nonlinearities. We will discuss various sources of these departures from the simplest model. Against this background, we can then consider two sources of individual differences in the acquisition of reading achievement. One source comprises differences in the parameters of the linear model: the estimates of the intercept (Y or X) and the slope for the individual student may differ from the average estimate over all students. A second source is a departure for the student from the bestfitting linear growth track: growth for a particular individual may not be steady over the years. Each of these sources of individual differences is a predictive challenge: of the various sources of data available about the student, which are able to explain the observed variation between students? The answer to this question is addressed in the second section of this volume.

The primary focus in this volume is on the measurement of growth in the several components of reading that are obtained from the

Interactive Reading Assessment System. Other dependent variables were included in the design as part of the year-to-year assessment; both standardized achievement tests and language data were collected from students, and teacher ratings were also gathered. In the last part of this section of the volume, we will discuss the application of the growth track approach to these other data sources.

Background

Research on student learning generally focuses on an omnibus measure of achievement (typically a standardized test score) at a single point in time (the posttest score) or at two points in time (both pretest and posttest scores). Neither of these approaches permits a trustworthy examination of the course of learning or development (Cronbach & Furby, 1970; Rogosa, Brandt, & Zimowsky, 1982; Rogosa & Willett, in press). Gathering data at a single point in time treats achievement as a static event; collecting data at two points in time is only a little better, because it does not permit any evaluation of the shape of the learning curve.

A foundational assumption of the Teaching Reading to Bilingual Children study is the notion that reading is a dynamic process, and that it was absolutely essential to tailor both the design of the measurement battery and the methods of data analysis to the character of changes in student performance over time -- more specifically, to the trends that occur over the several years that comprise primary reading instruction (kindergarten through fourth grade). As originally conceived, the study was to track students over the entire five years listed above. Regretfully, limitations in program funding meant that the original design was completed for only the first two cohorts. The remaining cohorts, which comprise the largest samples and cover significant population domains for purposes of generalizability, were tracked only through second grade. As a consequence, the first two cohorts are especially important for evaluating the adequacy of the analytic model for measuring growth, and for deciding on how to apply the model in determining significant correlates of growth.

Growth in Reading

Reading research (and educational research in general) has paid little attention in recent years to learning (Greeno, 1980). Questions like "What is the nature of the learning process?" or "What is the shape of the learning curve?" seldom arise. The emphasis, reflecting the burgeoning interest in the cognitive processes underlying skilled performance, has been on the manner in which an individual uses the mental resources available at a particular stage of development, without a concomitant concern about the way in which the individual acquired those resources.

Instructional programs include a developmental dimension, to be sure. Basal reading series are carefully graded to present the student with a sequence of materials and learning experiences that gradually increase in difficulty as the student moves through the series. (The discussion that follows holds for typical American reading series designed for teaching children to read in English. One can find distinct variations around the modal series, and programs for instruction in Spanish reading are generally quite different.) Readability is the criterion most consistently applied in the design of this sequence (Klare, 1974/75). As typically employed, this method consists in limiting the vocabulary that is available to teacher and student at a particular grade level. A secondary consideration is the length of sentences in instructional texts. This technique has been subject to serious criticisms by both researchers and practitioners (Kintsch and others), but practical alternatives have yet to be proposed and so the technique remains in use.

Viewed from this perspective, the dominant model of growth from a practical standpoint hinges on a steady change in the frequency of occurrence and letter/syllable length of the words used in instruction. In kindergarten and first grade, the student encounters commonplace words (the, of, she, said, cat, little, come, funny, children, and the like). These words present a mixed bag when viewed along multiple dimensions -- the decoding patterns vary widely, many of the spellings are irregular, function words predominate, and the high-information words that are often essential to writing are generally disallowed.

The primer materials used in the early texts of a given series begin by introducing a small number of words, typically two to four per lesson. These "new" words are discussed, used in a sentence, and pronounced (presumably to support rote acquisition of sight-word recognition). The meanings are already familiar to most students, and so the major consequence of instruction is sight-word recognition. The texts must be readable by students with limited decoding skills, and so are quite short, heavily "pictured," and generally incomprehensible as texts (for some reason, it is deemed inappropriate to incorporate longer and more coherent texts that could be read by the teacher for purposes of teaching comprehension).

Separate lessons are provided on phonic analysis of spelling patterns. The usual procedure is to begin with simple consonant correspondences using the short vowel sounds. By the middle of first grade, most children should have been exposed to consonant-vowel-consonant patterns.

The rate at which new words are introduced goes up during the primary years. In first grade, as many as 10-20 words may be presented each week; by second grade, counting the words that are formally introduced and those that appear incidentally (in text and in worksheets), the rate is increasing at about 10-20 percent per semester. Texts are longer and more complex, as are the sentences. In addition to the standard narratives that dominate in kindergarten and first grade, some expository passages begin to appear.

All in all, then, the course of learning represented by the typical reading series is a positively accelerated function, in which the several components of reading are combined in a single dimension.

After a slow beginning, in which students spend a considerable amount of time learning the routines of the reading lesson, while practicing materials with which they are already familiar, there is then a steadily increasing rate of introduction of new and more complex materials. By the beginning of fourth grade, relatively few new concepts are introduced through direct instruction. The changes are in the texts, which are presented for review of the content. It is assumed that, for practical purposes, the child has learned to read. Children with difficulty are recycled though a remedial program that covers the material just described, but at a slower pace, in smaller classes, and often in an individualized mode.

Designing an Assessment System to Measure Growth in Reading

What are the important considerations in designing a system for measuring the acquisition of reading skill and knowledge? In answering this question, we were guided in part by the separable-process model of reading. That is, we identified several components that we viewed as significant parts of the reading process, and that in principle might be independently assessed. In addition, we were guided by our analysis of the design of reading series. That is, we planned an assessment system that placed progressively greater vocabulary demands on the student.

Unique design features. The preceding considerations are not noticeably different from those that appear to undergird the development of standardized reading achievement tests. Our approach does include certain features that distinguish it from standardized tests. First, the tasks presented to the student were close to the demands of actual reading. Rather than assess decoding skills through a multiple-choice format which combines (implicit) decoding with other (implicit) skills, we asked the student to read a list of words (actually two lists). Rather than rely on a series of disjointed questions to determine comprehension of a passage, we asked the student to retell as much of the passage as could be remembered.

Second, the assessment of "reading" focused not only on the student's ability to handle printed materials, but equally on skills in handling the demands of formal language. Thus, several of the tasks required the child to respond to spoken rather than printed language, but in a context that placed formal demands on the response; for instance, students were asked to define selected words. By administering tasks that tapped in parallel response to printed and spoken language, we were also able to evaluate an assumption that seems to underlie the design of basal reading series: the readability limits incorporated in the series apply equally well to both forms of discourse.

Third, in measuring comprehension, we created texts at all levels of readability that met the criteria of coherence and comprehensibility. The passages were all designed to stand on their own merits as texts, without the need for interpretation in light of accompanying figures.

Finally, in addition to assessing performance on the various reading components, we also asked the students to explain how they approached each task. For instance, the student would be asked to articulate the word-attack strategy that he or she relied on in arriving at a correct pronunciation for a particular word, or to distinguish among various definitions that might be offered for a commonplace word.

The Interactive Reading Assessment System (IRAS). As noted in the previous volume, IRAS was designed as the primary instrument for measuring growth in the various components of the reading process. In this section we will describe how this system was fashioned to satisfy the criteria described above.

IRAS incorporates the developmental dimension of the basal reading series for each of the major components of the separable-process model -- decoding, vocabulary, and comprehension. In addition, each component is assessed in two or more ways. The test is interactive -- each student is individually assessed, and choices about the materials and tasks to be administered are based on the student's performance at the moment.

For example, the test begins when the student is presented a series of word lists graded by reference to several of the standard word counts used in preparing basal readers. The first list of words are those typically identified as appropriate to children in the first half of the first grade, the second list corresponds to words presented in the second half of first grade, and so on. The student is first asked for each of the lists in ascending order whether he or she can pronounce the words; when the youngster indicates that a limit has been reached, the tester asks the child to pronounce the words in the next easier list. After the actual performance limit is reached, then the student is asked to define words at that same limit. If successful, the student is moved up a list, and the task continues until an upper limit is found. The definition task focuses on vocabulary (i.e., word meaning) skills, and hence is administered orally.

Decoding is also assessed by a list of synthetic words. These lists are created according to analyses of the English spelling-sound system (Venezky, 1970), and are ordered according to several (non-frequency) factors known to affect difficulty of pronunciation. They also tend to proceed in the same order found in typical scope-and-sequence charts, although as noted earlier there is considerable variation in how different basals handle phonics instruction.

The texts used to assess comprehension increase over levels in overall length (number of words), propositional load (for practical purposes, this index equals the number of distinctive ideas; see Kintsch & van Dijk, 1978; Kintsch & Yarbrough, 1982), and text structure (Calfee & Curley, 1984; expository texts of increasing structural unfamiliarity were introduced from mid-first grade on).

IRAS materials were selected using word-frequency lists according to a linear progression in readability; to the degree that the basal materials drive the student's growth in reading, then a year of effective instruction should correspond to progress through two levels of IRAS. As noted above, success on Level A for each of the IRAS components should correspond more or less to the curriculum halfway through the first grade, success on level D should be expected for the youngster who has completed second grade, and so on.

To the degree that we have been successful in this approach to the selection of materials, IRAS scores provide a type of grade-equivalent information. The grade-equivalent measure derived from standardized tests has been subjected to several criticisms. Some of these apply to the IRAS indices, but others do not.

One problem with the grade-equivalent index has to do with the insensitivity at the extremes of a test. Most existing standardized tests are constructed to be "dense" within a relatively limited region of performance. A primary-grades test, for instance, is likely to contain items that are appropriate for students within a range of typical performance that covers first and second grade only. If the test is administered to typical third graders, most children will succeed on most of the items. The tests are generally scaled beyond the trustworthy limits, so that a difference of one or two items at the upper limits may lead to an increase of a grade level or more. IRAS was designed to cover a wide range of grades (from first through sixth) with equal sensitivity, using the interactive feature to quickly home in on those items most informative of the boundary between success and failure for a particular student.

A second problem with the grade-level index is more substantive than statistical, and has to do with overinterpretation of the index. A grade-level index of 3.0 suggests that the student should be capable of working with passages that are commonly assigned at the beginning of the third grade in a typical basal reading series. This interpretation is in fact reasonable, but the variability in materials also needs to be considered. First, there is considerable variability in the specifics from one basal series to another. The selection of words (more precisely, the constraints on the words available at a particular level of the series) is more or less the same, but one finds substantial variability in other facets of the texts -- the letter-sound correspondences, syntactic constructions, and the character of the passages. Second, the degree of consistency among series decreases over grades, largely because by third or fourth grade the constraints are negligible -- almost any word can be used. Secondary limitations (the number of polysyllabic words and sentence length) provide meaningful latitude together with word choice at these readability levels.

For all of these reasons, and also taking into account the variations in the character of phonics programs from one series to another, it is most meaningful to take the grade-level index as a measure of central tendency, rather than a fixed characteristic of a student or a series. In particular, it becomes important to examine not only the student's overall grade-level index (the usual approach with standardized tests), but to examine the profile of strengths and weaknesses across the various components of reading, a step that is possible with IRAS in ways that are not possible with the typical standardized test.

The preceding comments about basal series reflect existing practices for English-language materials in the United States. There appears to be less variation in the Spanish-language materials used in bilingual programs in the United States at present. This state of affairs arises in part because fewer series are available. In addition, there seems to be less in the way of fundamental variation; one might almost say that there is less "faddishness" in the design of these materials. This comment is not intended as a judgment on the merits of more or less variability, but rather to give our impressions. (Note: It would take considerable effort to carry out an appropriate comparison of the various materials, especially if this analysis were to be added to the student files. The basic information is part of the student file, and so the task is to examine the basal series with some care.)

IRAS: The critical level. The design of IRAS into components and levels for each component was coupled with an efficient but informative technique for determining the student's degree of proficiency for each component. In essence, the technique was to locate as quickly as possible the level at which the student first failed to meet a fairly lenient standard of performance. In general, as the difficulty of each task was increased, the student would do reasonably well for a while, and then there would be a relatively abrupt decline in performance. A lenient standard was set because we were interested in determining the furthest extent of the student's reach. If the student's response while decoding a word captured most of the letter-sound correspondences, we thought it important to establish this upper limit.

The details of how the critical level were determined are described in the IRAS manual (Calfee & Calfee, 1981; Calfee, Calfee, & Pena, 1979), and in Volume 5 of this report. An example will suffice for the reader who is unfamiliar with the instrument for the purposes of this volume.

The first task for the student was to scan a series of graded word lists, six words per list. The student was told that the words increased in difficulty from one list to the next, and was then asked to scan the lists until he or she encountered a list that was too difficult to "read out loud" (i.e., to decode). Virtually every student seemed to understand the task without apparent difficulty, and quickly went about searching the lists for his or her limit of mastery. At that point, the student was then asked to pronounce each word on the self-determined limit. If the student could correctly pronounce three of the six words, he or she was asked to try the next list. If the student did more poorly than this lenient criterion, the next easier list was presented. The process, which continued in one direction or the other until both a clear success and a clear failure had been obtained, almost takes more time to explain than the typical student required to perform it. Within a matter of two or three minutes, the tester had usually succeeded in determining the critical levels for the decoding of familiar words.

The critical-level method generated two pieces of information about each of the component tasks. One measure was the student's highest level of success, which in this report will be identified as the critical level. A second measure was an index of the quality of response at the critical level. A quality index was defined for analytic purposes as the average performance on this level.

The growth track. The design of IRAS, together with the technique for determining the student's level of competence on the instrument, lead to the postulation of a simple model of growth over time. The model, shown in Figure 1, represents change in student achievement over years of schooling as a straightline function. The correspondence with the grade-level index of the basal reading series is also displayed on the graph, along with the boundary limits for progress one year above and below the expected level. A typical student, based on the analysis of instructional materials that is incorporated in IRAS, should have trouble with the lowest level of IRAS in kindergarten, but should meet criterion (at least) on levels 2, 4, and 6 of the test when exiting from the first, second and third grades, respectively.

The normative model in Figure 1 can be clearly distinguished from the concept of grade-level used in standardized tests, and frequently excoriated in the literature. The authors of such tests rely on procedures similar to those used in the construction of IRAS to create a graded sequence of passages and related test items. After the items for a standardized test are written, a large group of students is tested, and statistical techniques are used to equate the levels of total test performance on the test to equivalent performance by students at a particular grade level. To repeat points made previously, criticisms of the grade-level index focus on the instability of the index at the extremes of a test (a valid comment), the degree to which the index can be generalized over variations in materials and classroom practices (also a valid concern), and the potential for misinterpretation (a problem with any summary measure). For the reasons discussed previously, we think that the IRAS criterion-level index is relatively resistant to these standard criticisms.

The grade-level index does have two strengths that make it attractive to practitioners. First, the index provides guidance for placement of individual students in basal materials that are reasonably appropriate to their decoding ability. An argument may be advanced that the placement may be misestimated, because a generalized placement is made based on one component of reading. While attentive to the problems consequent to such a decision, it appears nonetheless that this consideration is important in the attractiveness of the grade-level index to practitioners.

Second, an advantage of the grade-level index that is poorly articulated by most practitioners but is nevertheless of consequence,

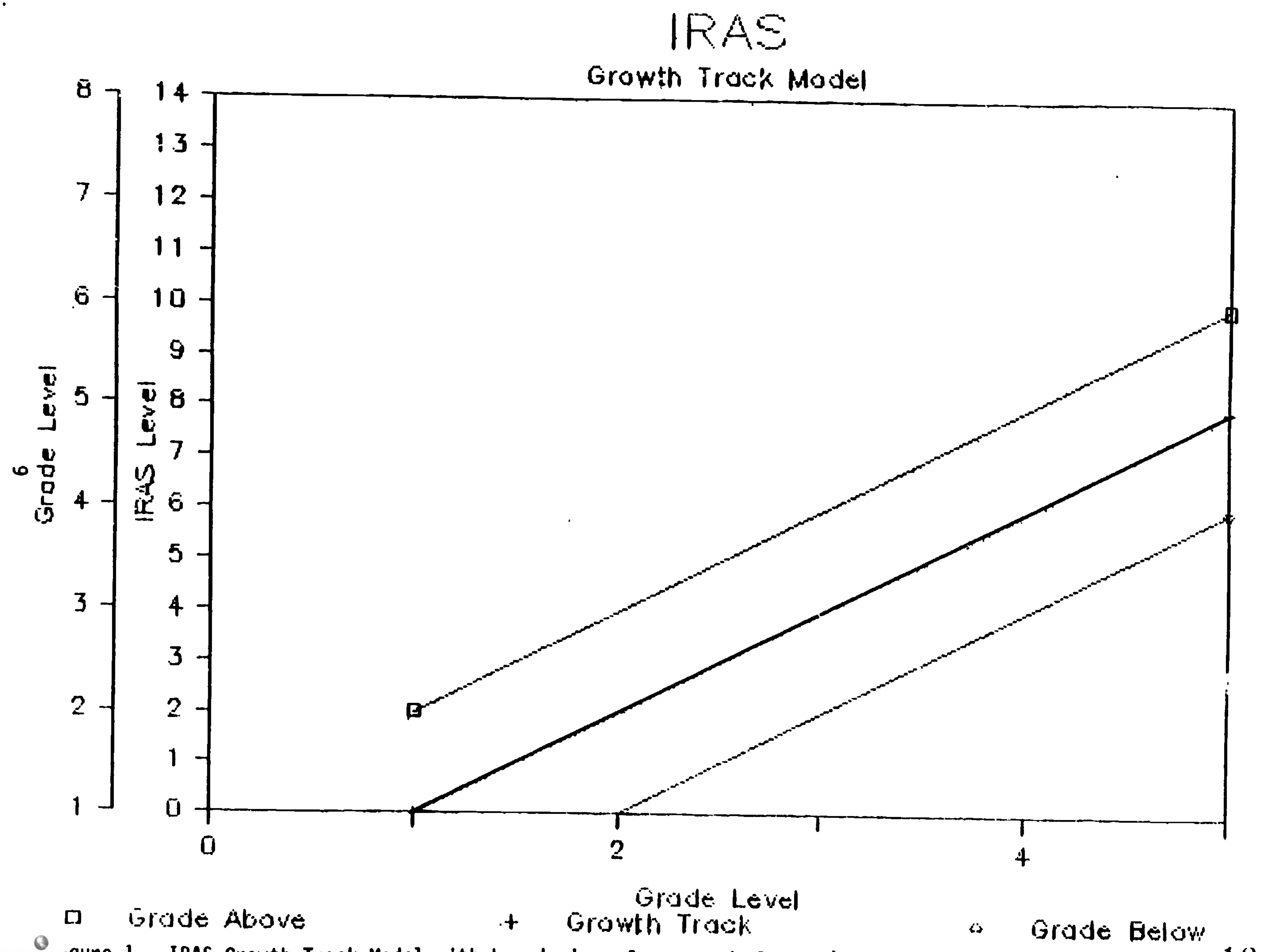


Figure 1. IRAS Growth Track Model with boundaries of one grade level.

is the imputation of a linear growth model. This feature appears when practitioners speak of "a year for a year" -- one year of progress on the grade-level scale for a year of schooling. The concept is simple, to be sure, but therein lies its strength. Standardized tests are flawed for this purpose, because the test constructors have not been able to attain adequate trustworthiness of the instrument at the extremes -- this problem has been solved in the design of IRAS, at least within the broad range over which the test extends.

At the risk of repeating caveats voiced earlier, three reservations should be stated with regard to the last comment. The design of IRAS is most secure at the primary levels. From Level 1 through Level 6, the fit of the IRAS materials to the typical basal series is reasonably good, in our opinion. From Level 7 upward, the amount of variability in prevailing practice is considerable, and provides less basis for grading materials. This comment does not imply that such variability is bad; to the contrary, we are inclined to think that curriculum materials from the fourth grade on should reflect the diversity found in the "real world."

Second, the fit of the IRAS design to existing basal materials is closest at the "word" level, and especially at the "sight word" level. Except for those programs that consist almost entirely of analytic phonics, youngsters are likely to be able to pronounce (decode) words according to the order of the graded series found in IRAS. For those reading components involving sentences and texts, the fit is likely to be less certain, which means that the linear growth model is likely to be compromised. Knowledge of word meaning (vocabulary) is a special problem. Basal series are constructed around the proposition that children do not know many words, and have to be instructed in the meanings of even the most commonplace words. We began with the suspicion that this assumption was wrong, so that the Definition data would not fit the projected growth track. Nonetheless, the model provides a basis for normative comparison.

Finally, it should be noted that IRAS is limited at the upper bounds. The most difficult words in the vocabulary series are quite demanding, both as regards pronunciation and meaning (mandatory, tumultuous, and veritable, among others). Nonetheless, one can imagine a more difficult set of words gauged either by pronounceability or familiarity. The Synthetic Word list is also relatively tough at the upper limits (e.g., euchormonium), but only a few such items were included in the test. In any event, some youngsters in the uppermost grades of the initial cohorts were able to perform quite well on the most difficult materials, evidence that IRAS is subject to a ceiling effect within the populations investigated. The ceiling is of greatest concern on spoken language tasks -- definition and listening comprehension.

Finally, we should point out two advantages inherent in IRAS, in its design and the accompanying methods of analysis, that are especially pertinent for a longitudinal investigation like this study. Two problems consistently arise in longitudinal research. First, it is not uncommon for the number of observations to vary from one individual to another -- students drop out or move, opportunities arise for an additional interview or observation. Second, intermediate data points are lost, leading to missed observations along a growth track. The consequence of these problems is that it can be misleading to present averages at different time points; different subjects are represented in the averages, making comparisons difficult to interpret unless the number of subjects is quite large and the proportion of missing observations quite small. In this report, we will present time-point averages on occasion, generally with a word of caution when the averages are not comparable. For most of the analyses, we will rely on the estimates from the growth-track model, which have the advantage that all students are equally often represented at each point in time, based on the best-fit estimates from the linear model. There is a price to be paid from this strategy -- we are working with estimates rather than actual data points. The purpose of the following section is to present data from the various cohorts that demonstrate the extent to which these estimates are likely to be trustworthy as measures of general performance trends.

Analysis of Average Performance

Most of the IRAS measures were designed to follow a linear growth track; certain exceptions to this generalization will be noted later in this section. Ideally, the study would have tracked all students for three or more years, permitting a clear test of this hypothesis. Students in the first two cohorts were actually tested from kindergarten through third or fourth grade, and so nonlinearities can be assessed for individual youngsters. Students in the third and fourth cohorts, which represented the largest and most representative segment of the sample, were only tested on reading achievement during first and second grade; with only two data points, the linear model cannot be evaluated.

In this section, the average performance on the various IRAS measures will be considered as a function of the number of data points available per student. Two questions are of primary importance in this analysis in establishing the foundation for the later analyses of IRAS reading achievement. The first question is: to what extent do the data follow a linear progression for those cohorts with three or more data points? The second question is: to what extent are there differences in the year-to-year averages of groups with different numbers of data points?

If it appears that the predicted linearity is observed, supporting the proposed growth track model, then we will be justified in summarizing each student's performance by estimating the intercept level on entry to first grade and the slope or rate of growth from that point onward. This strategy is especially supportable if the year-to-year differences in average performance appear slight over groups with different numbers of data points. In contrast, problems in this strategy may arise from either of two sources. First, we may find evidence of nonlinearities where the data allow such a test, in which case we will have a more difficult time in comparing groups with differing numbers of data points. Second, it may appear that the various cohorts perform u ferently durin, those years when they can be compared; this problem is less serious than the first, but would be troublesome if it were to appear.

In the remainder of this section, we will examine in some detail the data pertinent to an evaluation of the growth track concept. An overview of the findings may be useful as a road map through the detail that follows. Two points summarize the most important results. First, the growth track hypothesis receives strong support from the average performance levels for all of the IRAS measures to which this hypothesis is pertinent (as will be noted later, spelling and sentence reading are measured on different scales, and linear growth is not predicted). Moreover, the average performance levels are quite close in certain critical instances to the absolute levels predicted from th design of IRAS. Second, for children from more or less comparable b ngual backgrounds, performance during a given year is virtually entical from one cohort to another, making it reasonable to compare data from the different cohorts. This comparability needs to be leavened with one caveat. Those cohorts with only two data points, and those in the early grades, are subject to the greatest variability in projections to performance in later grades. For example, when we use the growth track model to predict achievement levels at the end of fourth grade for all groups, as the model permits us to do, the estimates will be most trustworthy for those students for whom we have data from first through fourth grade, and least trustworthy for students who were tested in first and second grade only.

We will begin the presentation of these findings with a detailed accounting of the Vocabulary Decoding (VDC) measure from the English IRAS. This discussion will serve as an illustration of the procedures and the data structure of this instrument. Afterwards, we 1 then present the remaining English IRAS measures in less detail fashion, followed by the Spanish IRAS measures. To repeat a point made earlier, the ief purpose of this section is not to consider the substantive import of the findings from reading achievement measures, but rather to establish the feasibility of using growth track estimates to allow analyses across cohorts with differing numbers of observations. From time to time, substantive comments will be interjected as appropriate, but the bulk of the analytic work on IRAS will appear in Volume 5.

The Growth Track of English Vocabulary Decoding (VDC)

Basic descriptive statistics. In Table 1 are the basic descriptive findings for the IRAS VDC measure. On this test, the student was presented a series of word lists graded on vocabulary. An interactive search strategy was followed to determine the level at which the student passed a lenient criterion (half the words properly pronounced), but failed the criterion on the next most difficult list. The word lists corresponded to half-grade increments; success on the first list was equivalent to words used through the first half of first grade, success on the second list was equivalent to words used through the second half of first grade, and so on.

Table 1

Interactive Reading Assessment System - English:
Descriptive Statistics for the Vocabulary Decoding
Scale for Individual Cohorts

Bilingual Cohort	Statistic	Year 1	Year 2	Year 3	Year 4	Y-Intl	Slope
2-years	M	2.01	5.44			-0.39	3.43
	S	2.40	3.25			2.92	2.27
	N	152	152			152	152
3-years	M	1.92	5.42	7.78		0.06	2.91
	S	2.20	4.02	4.29		2.42	1.61
	N	38	37	37		38	38
4-years	M	2.94	4.77	8.07	10.10	1.02	2.48
	S	2.47	3.50	4.47	4.12	2.75	1.03
	N	56	56	56	56	56	56
Monolingual English							
2-years	M	4.63	6.94			3.02	2.31
	S	2.80	2.70			3.98	2.53
	N	36	36			36	36

Note: Tabled values are based on critical indices where one IRAS unit equals .5 grade-levels.

The table is organized along the left margin according to cohorts -- bilingual students who were tested during first and second grade, those tested for grades 1 to 3, and those tested during grades 1 to 4, followed by the monolingual English cohort which was tested only during the first two grades. (The identification of the various cohorts as two, three, or four years is somewhat misleading. Most of the students actually entered the study as kindergartners, and so the number of years is actually three, four or five. The purposes of the present discussion, which focuses on reading achievement measures, seemed better served by identifying the cohorts according to the number of years during which reading achievement was actually measured).

Three statistics (mean, standard deviation, and number of observations) are shown for each of four instructional years. Statistics are available for all cohorts for the first two instructional years; there are no data for the third and/or fourth instructional years for the cohorts who were only followed for two or three years. In addition, estimates of the Y-intercept (at first grade) and the slope are shown in the two righthand columns. These two sets of statistics will be discussed following presentation of the year-to-year measures.

Year-to-year averages. To get a sense of the meaning of the year-to-year data, let us consider some of the data for Instructional Year 1. The students in the final cohort were tested only during first and second grade. This cohort, the largest in number, contains 152 students. Their performance on the VDC measure at the end of first grade was 2.01; the children could on the average decode words in lists of readability corresponding to end-of-first grade. There was considerable variability around that average level, ranging in IRAS levels from 0.0 (failure to meet criterion on the easiest list) to 6.0 (success on materials commonly found at the end of third grade).

The data for the 3-year cohort, a smaller sample with 38 students, look quite similar to that of the 2-year cohort -- the children read on the average words of a difficulty level appropriate to the materials in the basal reader to which they are likely to have been assigned. The 4-year students, who comprise a sample of 56, reached an average IRAS level of almost 3.0, corresponding to mid-second-grade difficulty. The monolingual-English sample, 36 students who were part of a two-year cohort, could decode real words considerably higher on the readability scale than the bilingual groups described earlier. The average IRAS level of 4.6 is equivalent to words found at the beginning of third grade in the typical basal reader.

The remaining year-by-year data in the table can be briefly characterized as follows. First, all of the groups made steady progress on decoding real words over the years. The averages varied somewhat, but in general the increases were about two IRAS units -- roughly equivalent on a readability scale to one year of growth for each year of instruction. Second, the variability in individual performance increased over the years. This increase is partly artifactual; there is a "floor" of zero, below which performance cannot be assessed. The test also has an upper limit of 14.99 (roughly equivalent to seventh grade performance). The extent of variability is nonetheless notable to the eye -- after first grade, relatively few of the bilingual students were at either the floor or the ceiling, but they were widely dispersed between the two limits (some monolingual English and Spanish students did reach the upper level of this test when tested at the end of second grade, but the number who topped out was not excessive).

Graph of the year-to-year data. In Figure 2, the data from Table 1 have been plotted along with the predicted growth track. As can be seen, for the bilingual cohorts, the averages fall close to the growth track on exit from first grade, show an increase to a grade level above the growth track at the end of second grade, and then steady growth which parallels the growth track in the remaining years. Moreover, for these cohorts, there is little noticeable difference from one cohort to another. Some of the variations, though slight, have implications that are not immediately obvious but are of significance nonetheless. The monolingual English students also show growth which parallels the growth track, although they are substantially above it.

For instance, the 2-year cohort made a gain of almost 3.5 IRAS levels from the first to the second grade. No other group made an average yearly gain of this magnitude. The 3-year cohort made an equally substantial gain from first to second grade, but progress was slower from second to third grade. Since testing of the 2-year cohort had to be interrupted at the end of second grade, we have a less trustworthy measure for this cohort over longer time spans. The data in Figure 2 warrant the use of the growth track model for analysis, but care must be taken in interpreting differences from one group to another when the estimates of slope and intercept are based on different numbers of observations.

The data in the figure do lend themselves to a fairly simple interpretation. The bilingual students enter school with little knowledge of English print. It appears that they are not taught much about print in kindergarten. Beginning in first grade, the basal reading series is used for instruction; progress through the series is dictated by the readability index. By the end of first grade, these students have been exposed to words corresponding to the second level of IRAS; they can read words of this difficulty, at least on the average. Progress continues in linear fashion over the time course of the assessments. We suspect that students have little experience with English print outside of the school setting, and so they are learning what they are taught within that setting. The data for the monolingual English students are more conjectural, but it appears that they are more familiar with printed English on arrival to school and/or receive more exposure during kindergarten. In any event, the data in the figure suggest that these students have an advantage when they enter first grade, but do not appear to learn at a faster rate, at least not as performance is measured by the VDC index.

These brief interpretive remarks are intended only to give the reader some sense of the "meaning" of the rather abstract representations in the table and the figure. In later sections we will give

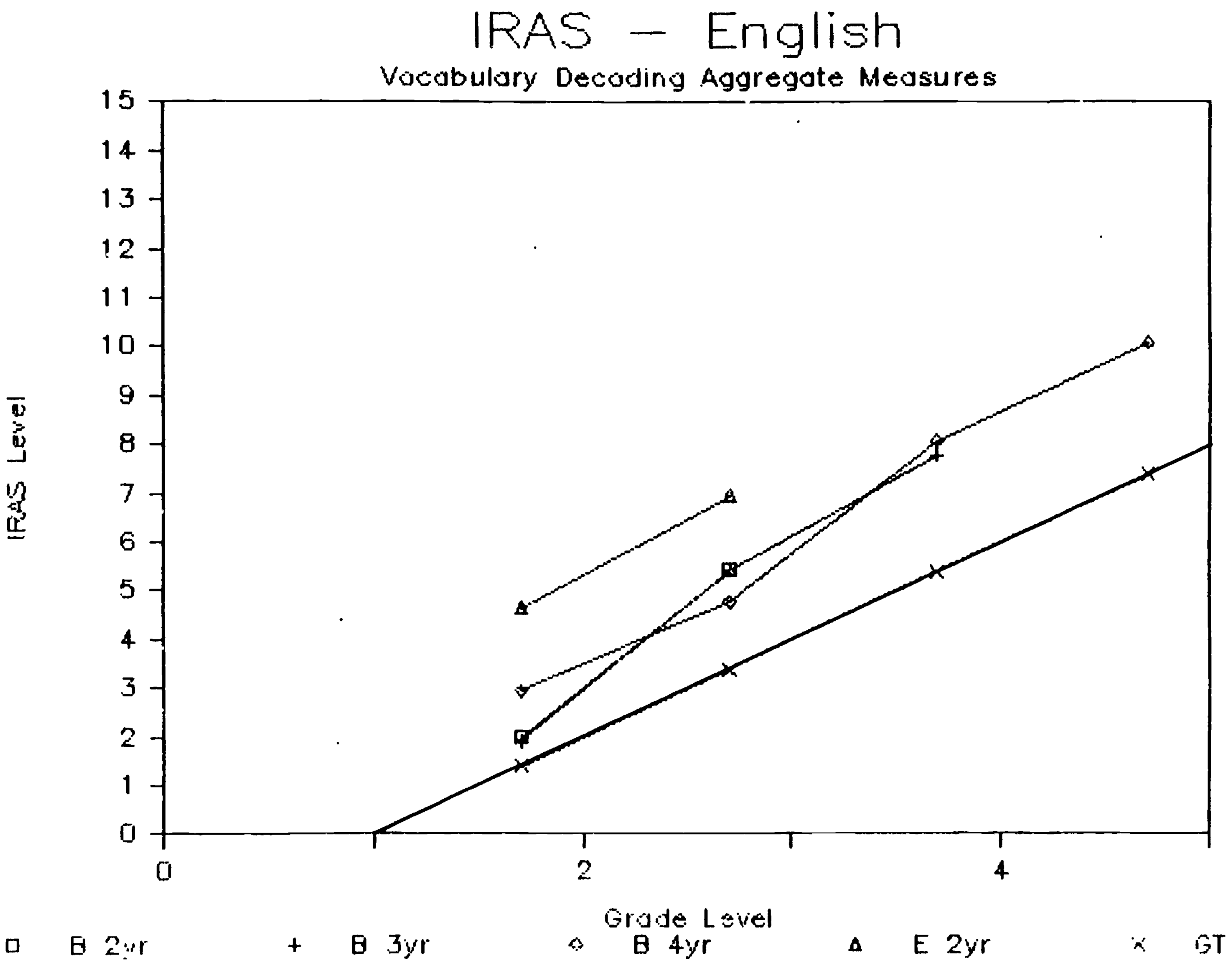


Figure 2. IRAS-E Vocabulary Decoding aggregated values for individual cohorts.

closer attention to the issues raised by these data -- what does entry-level performance look like; what is the character of the kindergarten and primary instructional programs; and so on? For now major points to be made are that (a) performance on the VDC measure does appear to follow the growth track postulated in the design of IRAS, and (b) most of the cohorts show little substantial difference from one another. These findings seem to us to warrant the use of the growth track as an analytic tool for estimating (a) the individual student's level of performance at some predetermined entry point, and (b) the individual student's rate of growth in response to instruction.

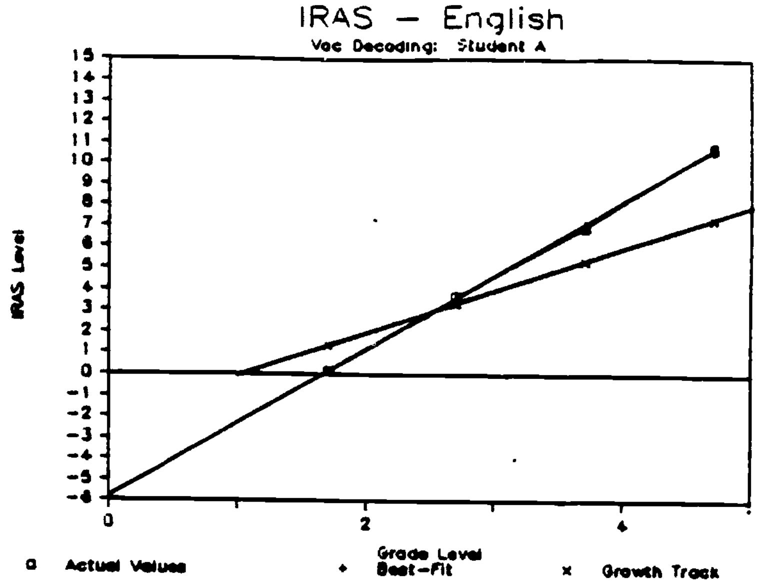
Estimating the parameters of the growth track. Figure 3 illustrates the techniques of parameter estimation by means of the growth track model. Three students have been selected for purposes of illustration to show the range of variation in performance that is possible within the limits of the model. Some cases do not fit the model; these will be discussed later in this section.

The top graph, for student A, shows the IRAS levels along the left axis, and grade in school on the bottom margin. The growth track is the 'X' line moving from the lower left to the upper right corner of the graph. The track begins at a level of 0.0 for entry to first grade, then moves in a straight line direction, two IRAS levels for each year of schooling, to a value of 8.0 on exit from fourth grade.

Student A did not pass the first level of the VDC measure at the end of first grade. (IRAS was administered in March of the school year, and so the measures are plotted not at the end of the year, but at a value seven-tenths through the year.) The observed score of .17 means that the student managed to read aloud one of the six words. By the end of second grade, the student had made considerable progress; with a score of 3.72 on the VDC test, the student was at the expected level. Progress continued to be excellent during the next two years (this student is obviously from the first cohort, which was tested in grades one through four), and the student was more than a grade level above expectation at exit from fourth grade.

The progress of this student, while not falling on the expected growth track, clearly fits the linear growth model. The rate of growth is a steady 3.5 levels per year, almost twice the rate expected from the design of the basal series. One can apply a ruler to the data points and determine the slope and intercept of the best-fitting linear function without resorting to more precise statistical methods. The latter (a straightforward application of the linear regression technique) was used to estimate the slope and intercept for the linear growth model for each student in the study.

While the rate of growth index can be estimated without ambiguity from a set of observations like those for student A, the estimate of the intercept requires some thought. The question has nothing to do with statistical procedures, but with the matter of deciding on the point from which progress should be measured. What is the appropriate starting point from which to measure growth? From a mathematical



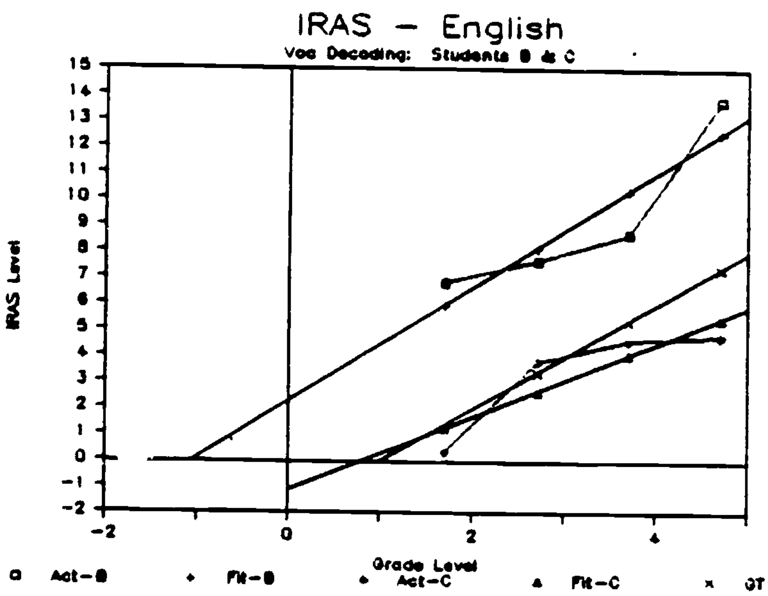


Figure 3. IRAS-E Vocabulary Decoding actual values and best-fit regression line for students A, B, and C.

perspective, one might chose the "zero-point" on the year-in-school scale; zero on this scale corresponds to entry into kindergarten. A better answer, it seems to us, is to set the entry point at the tim when reading instruction begins. A naive point of view, but justifiable in our opinion, is that for many of the children in the study systematic experiences with printed English were unlikely to have taken place prior to the beginning of first grade, and so this time should be chosen as the starting point for growth. A variation on this reasoning, one that will be explored in greater depth later in this volume, holds that one should determine the actual starting point of instruction -- before kindergarten entry, during kindergarten, during first grade, or perhaps at a later grade -- and measure growth from that time. For the present, we will use first-grade entry for determining the estimate of the student's entry level.

For student A, the intercept estimate (at first grade) is negative, -2.3, implying perhaps that the student was lacking in some of the prerequisite skills for beginning to read. We will suggest another interpretation below. For now, the estimate can be taken as a mathematical abstraction reflecting the fact that at the end of first grade, student A was performing substantially below the level expected from an examination of the first grade basal materials. Growth from that point on was strong and steady. Because the growth is linear, we can summarize student A's performance by two summary estimates, the intercept and the slope, which we can interpret in turn as the student's achievement level at the beginning of first grade and the average rate of growth from that point onward.

The bottom graph in Figure 3 shows the data for two other students, whose patterns of growth are markedly different from student A. Both students are from the 4-year cohort. While both show clearcut growth over the four years, the changes are much more erratic than in the first example. The slope and intercept estimates capture significant features of the growth patterns, but with less fidelity than was the case for student A.

Student B does extraordinarily well on the VDC index when initially tested at the end of first grade; performance is at a level expected of students finishing third grade. Progress then is slow for the next two years, followed by a burst during third grade. For practical purposes, this student had reached the top of the test by the end of third grade.

The best-fitting linear fit for this student is shown as the '+' line in the figure. The average growth rate, 2.21 IRAS units per year, parallels the rate of the expected growth track. The intercept on entry to first grade, 4.42, suggests that the student may have had some initial exposure to printed English in kindergarten or before. This hypothesis could be checked by examining measures of kindergarten performance for the child. It is also possible that the student was in an especially effective first-grade classroom; the observational data would be of value in testing this hypothesis.

The third youngster used to illustrate the methodology performs almost identically with student A when tested at the end of first and second grade, after which there is little or no further growth. The C protocol, like that of student A, contains a significant amount of nonlinearity. The slope of the best-fitting line nonetheless provides a good estimate of the average rate of growth over the four test points (about 1.4 IRAS units per year, which is .6 units below expected). The intercept, swung around as though at the end of a rope (more properly, at the end of a seesaw), is more significantly affected. Rather than having a negative value as for student A, the intercept is moved upward to a level fairly close to zero at first-grade entry.

Both the B and C protocols illustrate an important point: although the average levels of performance for groups of students may trace out straight lines, this finding does not mean that the paths for individual students are necessarily straight. Volume 5 will report more detail on this matter. For now, a couple of summary remarks are offered. First, data for the typical student are reasonably close to the linear model, and so the estimates from the model do a reasonably good job of summarizing progress. Second, the deviations from linearity may be partly due to "noise" in the data, and partly due to floor and ceiling artifacts -- we have taken considerable care to examine these possibilities. But we also suspect that some of the deviations reflect the effects of year-to-year instructional variations -- these sources of nonlinearity will be the focus of discussion later in the volume.

Average slope and intercept functions. For most of the analyses reported in the later volumes of this report, we will rely on the slope-intercept estimates to summarize patterns of student achievement. One way to track growth over time is to test each student in a cohort at each of several points in time, and then display the averages for each time point. This approach has a number of shortcomings.

Some of these limitations can reflect limitations in the data. If every student is tested an equal number of times, and if there are no missing data, then at least the averages are all based on the same set of individuals. It is not uncommon, when one is working in applied settings, to find these desired features of a data structure honored in the breech. Such was the case with our study. As noted already, funding limitations meant that some students were tested over a longer time course than others. Moreover, despite strong efforts to ensure that every student in the sample was tested on a predetermined schedule, some observations were lost. While the percentage of missing data is remarkably small in this study, nonetheless there are some blank spots in the longitudinal record for a few students.

Other limitations arise from the inability of simple averages to capture patterns of individual performance. Consider the averages displayed in Figure 2. This plot shows a remarkable congruence with a predicted straightline function. Yet this display tells us nothing about the way in which individual students perform over time. As Estes (1956) observed some years ago (as have others before and since), aggregating data into means and variances entails a loss of information about the character of the original data structure, and can easily lead to misinterpretation of the character of that structure. As the B and C protocols indicate, the straightline averages in Figure 2 represent in part student growth patterns that are noticeably nonlinear -- as it turns out, although these three protocols were selected almost at random, if you compute the averages at each year, the result is very nearly a straight line, even though one function is straight, one bent upward, and the third bent downward!

Summarizing each student's growth over years by the slope and intercept (and by also calculating a measure of nonlinearity), we have been able to handle the problems mentioned above. Missing data points reduce the sensitivity of the estimates, but do not otherwise bias them. By calculating the mean and variance of the slope-intercept estimates, one can determine rather precisely the character of individual difference in growth patterns, and can separate effects that are related to entry level from those that are related to growth.

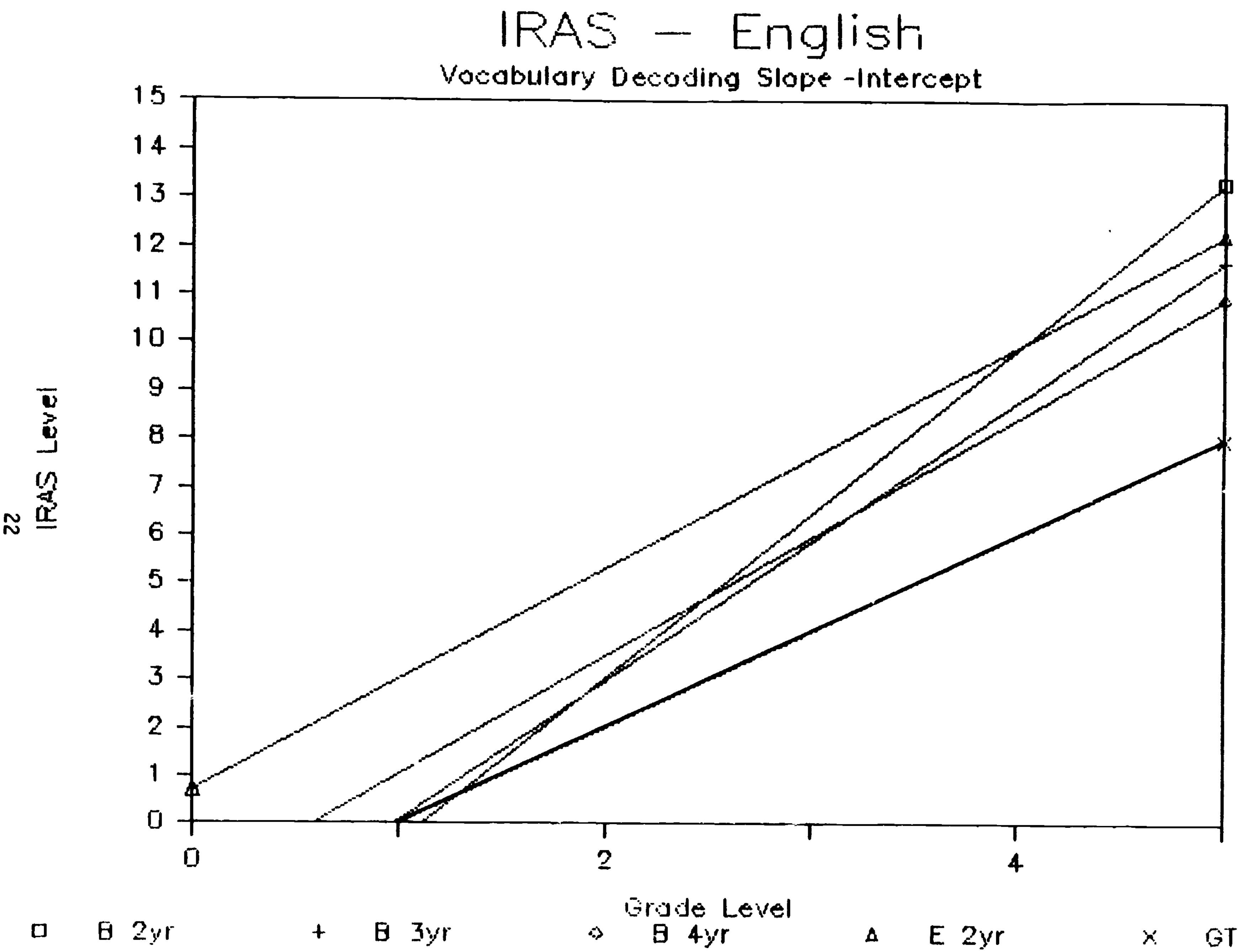
One can also use the slope-intercept estimates to compare different groups, even in the presence of unequal numbers of time points or missing observations. In Figure 4 we show the same data presented earlier as Figure 2, but using the slope-intercept estimates as the basis for representing the data. The effect of using the estimates is to smooth the minor fluctuations that occur from year to year, and to allow extension of the 2- and 3-year cohorts over the full range from entry to first grade through exit from fourth grade. Please remember that in making this extension, predictions at the later grades are subject to more unreliability for those cohorts with fewer time points.

The slope-intercept picture is quite close to the original year-to-year averages for most cohorts, not surprisingly given the linear trends in the averages. The slope-intercept projections for the 3- and 4-year cohorts parallel the expected growth track. The 2-year performance is noticeably higher at the later grades; the slight departure in the direction of a higher slope from first to second grade is amplified at the later grades, and should be taken with a grain of salt. Nonetheless, the degree of correspondence is quite good, and warrants the application of the growth-track methodology for examining the various longitudinal measures.

Growth Tracks for the IRAS Measures

Having illustrated the basic techniques with the English IRAS index of Vocabulary Decoding, we will now examine the other IRAS measures, first English and then Spanish. For each index, we will present the year-to-year averages for each cohort, followed by the results from the slope-intercept method.

English IRAS. The Interactive Reading Assessment System contains a number of subtests, which in combination cover the major domains of reading -- decoding, vocabulary, and comprehension -- from both the perspective of oral reading and formal command of spoken language. The nine subtest indices will be organized into three clusters reflecting



22

148

Figure 4. IRAS-E Vocabulary Decoding slope-intercept values for individual cohorts. 149

substantive relations rather than the order in which the subtests were administered, the latter being based on convenience and consideration for the student and the tester. We will rely on graphs of the data for the basic presentation; complete statistics are tabled in Appendix A.

The first cluster includes the subtests of Vocabulary Decoding (VDC, discussed previously), Letter-Sound Correspondences (LSC), and Letter-Sound Spelling (LSP). These three subtests place primary demands on the students ability to pronounce or to spell isolated words, either real or synthetic.

The second cluster, also comprising three subtests, includes Vocabulary Definition (VDF), Narrative Listening Comprehension (NLC), and Expository Listening Comprehension (ELC). In each of these subtests, the interaction with the student is entirely through oral communication; no decoding is required. The student is asked in turn to define a word, or after listening to a either a narrative or expository passage, to retell as much as can be remembered. The emphasis in all of these subtests is the level of skill in formal language tasks.

The final cluster of three includes Sentence Reading (SRD), Narrative Reading Comprehension (NRC), and Expository Reading Comprehension (ERC). These tasks come closest to the commonplace meaning of "reading." In Sentence Reading, the student was given a graded series of sentence sets, each taking less than half a minute to read. The child read through the series until a time limit was exceeded; the time limit was set for each series so that increasing fluency was required for successively more demanding sentences. In Narrative and Expository Reading Comprehension, the youngster read one or more passages, either aloud or silently, depending on the level of acquisition, and then recalled as much as possible. The emphasis in the task was on comprehension, not decoding. However, in order for the child to have any chance at comprehension, some degree of fluency in decoding was essential.

The upper panel in Figure 5 should be quite familiar if you have read the preceding section of this volume, since it simply recapitulates the data from the VDC subtest. The basic conclusion from these aggregate data is that they provide strong support for the growth track model.

Immediately below in the middle panel are the data from the LSC subtest. Because of the varied arrangements in the decoding scope-and-sequence charts of most basal series, it is impossible to project a growth track. The LSC subtest comprised six levels. The first lists, A and B, assessed the most basic consonant-vowel correspondences, including the long-short vowel contrast. Levels C and D tested more complex Anglo-Saxon patterns, including consonant and vowel digraphs. Finally, the Romance and Greek spelling patterns that students must deal with from about third grade onward were presented in levels E and F. Accordingly, we would project a more or less linear progression over the grades on the LSC index, with a limit close to the upper boundary of the index by the end of fourth grade (implying a growth

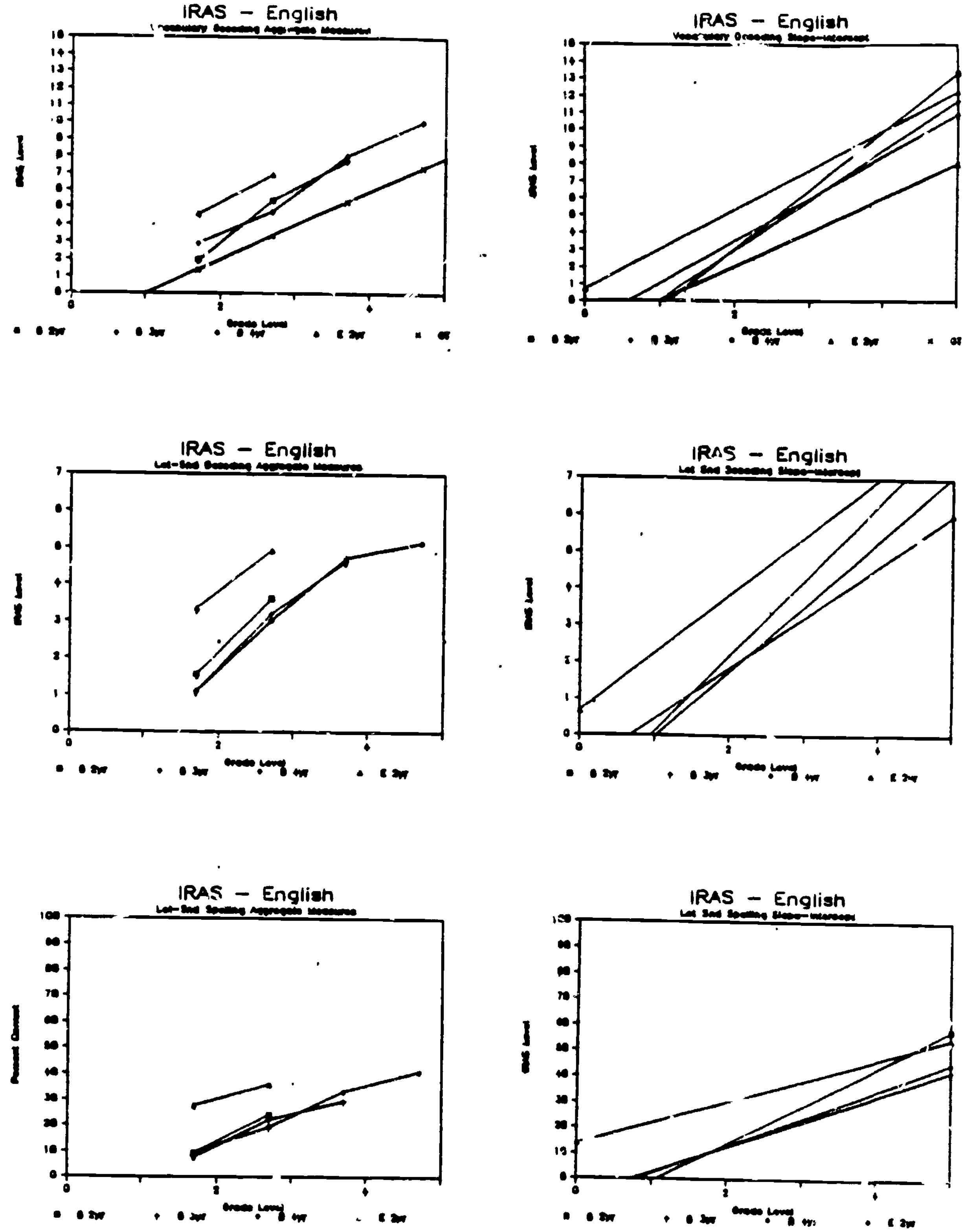


Figure 5. Growth plots of IRAS-E decoding scale indices for aggregate measures (left) and best-fit regression lines (right) for individual cohorts.

rate of about 1.5 IRAS units per year), but with considerable variability at each level.

The observed data in the lefthand middle panel are actually rather intriguing. The progress during the primary grades is steady and about on the mark that we would predict from our examination of basal series. Students learn the basic correspondences from the Anglo-Saxon layer of the language (Venezky, 1970; Calfee, 1982), which are subjected to intensive worksheet practice in most basal series. Most reading programs tend to deemphasize decoding from fourth grade onwards, which means that the Latinate and Greek spellings receive relatively little attention.

A second feature of the LSC data is the tendency for a systematic curvilinearity in performance over the entire span of grades. The greatest increase is from the end of first grade to the end of second grade, and the subsequent increments are smaller. Accordingly, the linear estimates of progress shown in the righthand panel favor the cohorts that were tested the fewest times -- the estimates for the 2-year cohort show the highest growth rate, and the 4-year estimates are lowest. These estimates are probably biased because of the systematic nonlinearity in progress in decoding. The monolingual English growth estimate is probably untrustworthy for the same reason, although it is clear that this cohort has a better grasp of the abstractions of decoding at the end of first grade than do the students in the bilingual cohorts.

The Spelling subtest was patterned after the LSC materials, but the scoring of the task was different. Students were asked to spell a list of synthetic words. The index, chosen for reasons of convenience, was the percentage of correct spellings. Again, the shape of the growth track function cannot be defined a priori; we would expect steady progress, perhaps increasing in the early grades, and then slowing down as the students began to reach a level of mastery.

In fact, most of the students did rather poorly on the Spelling subtest. As shown in the bottom panel of the figure, there is steady progress, but the rate is slow. By the end of fourth grade, the average student could correctly spell less than half the items on the list of synthetic words, and the rate of progress is slowing down. The monolingual English cohort was at an advantage when tested at the end of first grade, but the estimated rate of progress suggests that they would be roughly equal to the bilingual cohorts on entry to fifth grade.

Decoding is one of the most serious hurdles to challenge the beginning reader. While mastery of decoding skills is important, the student must also learn the "technology of language" in order to become fully literate. In Figure 6 are shown the IRAS data on three indices of formal language competence, where decoding skills are eliminated as a barrier to performance.

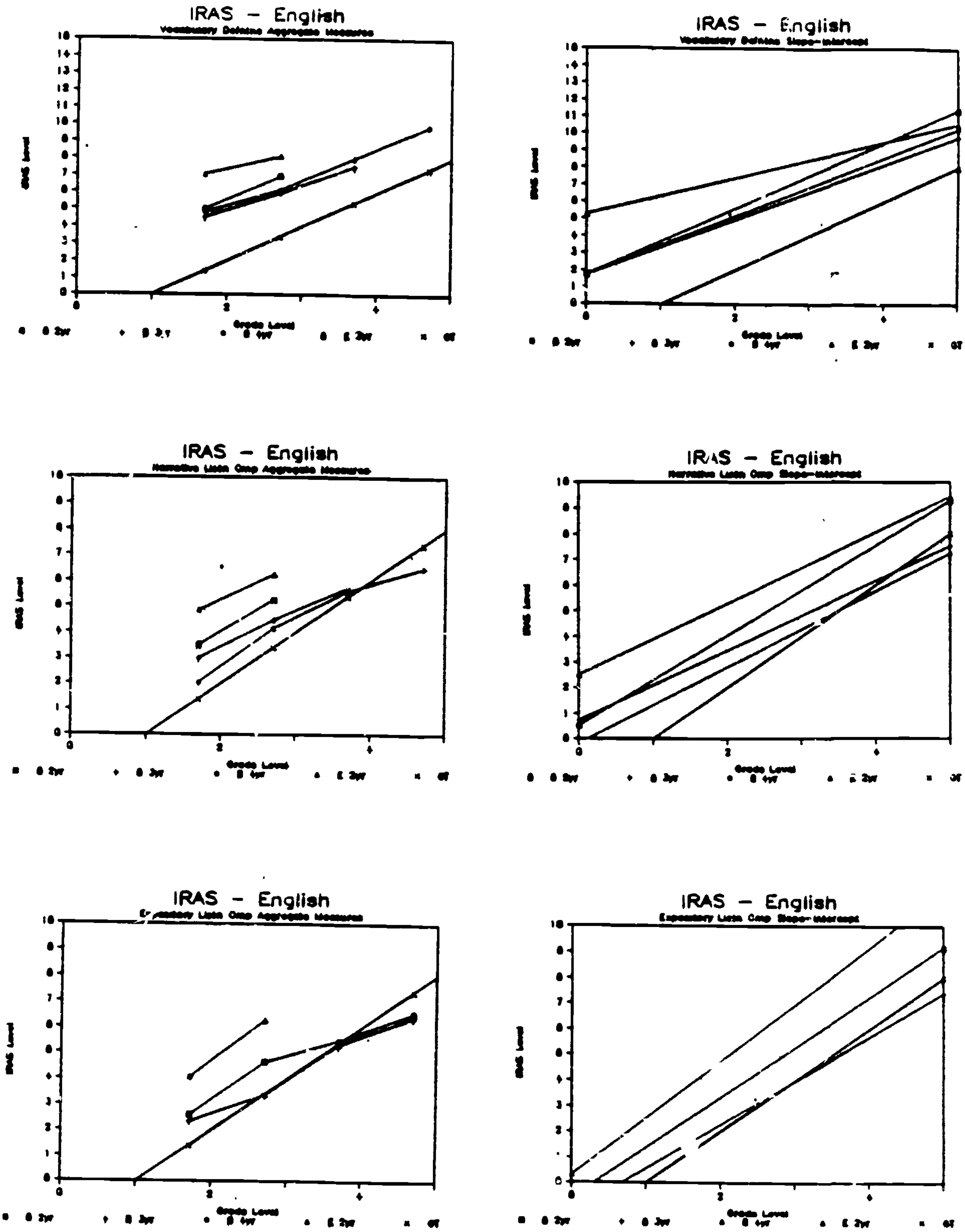


Figure 6. Growth plots of IRAS-E formal language scale indices for aggregate measures (left) and best-fit regression lines (right) for individual cohorts.

The upper panel of the figure shows achievement levels on the subtest of Vocabulary Definitions. The student was asked to define a word -- what does it mean; can you think of another word that means about the same thing; does the word mean X, Y, or Z (a multiple-choice probe)? If the student managed to demonstrate an understanding of half of the words in a list by any of these criteria, then the student was given credit for performance at that level of IRAS.

The students in the bilingual cohorts performed remarkably well on the VDF subtest; to be sure, the monolingual English cohort achieved an even higher level of performance, but to focus on this difference is to miss an important point. If we take as given the validity of the word-frequency counts on which readability indices are based, then the typical student in the bilingual cohorts understand the meanings of words at about the third-grade level when they leave first grade. To put it in a different perspective, these students enter first grade possessing an oral vocabulary corresponding to basal materials designed for students beginning third grade. In other words, the average bilingual student in our sample had a command of English that was not to be challenged for two years.

There is some evidence of learning. The year-to-year increase is noticeably less than the predicted rate, roughly half what one might expect from word frequency counts. The data from the 2-year monolingual English cohort is also interesting -- they begin at a clear advantage over the bilingual cohorts, but if the linear growth track is valid, then by the late elementary grades one should find a convergence in the performance of these two groups of students <u>if vocabulary performance is tested in the manner used in IRAS</u>.

Finally, performance on the VDF subtest does seem to fit the linear growth track model. To be sure, the slope and intercept of the average performance plot do not follow the standard readability predictions, but we think this discrepancy can be explained. The point is that change in average student performance is linear over the range of the study for each of the cohorts. The plot of slope-intercept data in the righthand panel closely mirrors the aggregate measures to the left, further supporting the growth-track model.

If the findings for the VDF subtest have general validity, one might be concerned about the educational implications. Such an interpretation needs to be portrayed against a framework that includes information about entry-level indicators and the instructional program. Nonetheless, at first glance it appears that students in all cohorts possess a command of English vocabulary considerably in advance of the demands posed by the basal series. To put the matter more directly, it appears that most of the students are asked to spend time during the primary grades studying words that are already part of their working vocabulary.

The two lower panels in the figure suggest that the pattern of performance in VDF is similar to that for the "listening comprehension" of narrative and expository passages. A fundamental difference does exist between the narrative and expository graphs, to be sure. Moreover, definition and comprehension also follow different courses.

In both the comprehension tasks, the student was asked to recall the passage; the basic comprehension question was posed to the student -- "Tell me what you just read." Probe questions were also asked of the student in order to determine whether the students had knowledge that was not readily rememberable.

Let us consider first the Narrative performance. The growth track model is applicable here, in principle. When tested at the end of first grade, the bilingual cohorts are on the average above the readability level predicted by the growth track model. Year-to-year growth is approximately linear, but with some indication of a slowing of the rate of progress. Compared with the VDC profile, the picture is one in which bilingual students in the study can handle the narrative texts presented in the first three grades, but their competence in handling connected texts appears to decline over the years of schooling, when compared with the expected level based on the readability standard.

Moreover, the comprehension of narrative texts is substantially below the level of competence reflected in the VDF index. For instance, when tested on narratives, ending second-grade students from the bilingual cohorts can comprehend connected text at the level considered appropriate to their grade level. However, according to the VDF scale, where both assessments require only oral command of formal language, these cohorts can handle third-grade words. Vocabulary is presumed to drive comprehension, if we are to believe the readability experts, but from the IRAS protocols, it appears that comprehension lags behind vocabulary in the performance of bilingual cohorts.

Finally, with regard to narrative performance, it should be noted that the monolingual English cohort once again shows a tendency to converge (based on the slope-intercept estimate) with the bilingual students. It is as though the monolingual English students in our study begin with an advantage on entry to school, but become more similar to the bilingual cohort when tested on oral comprehension of narratives as they entered the upper elementary grades.

The Expository Listening Comprehension (ELC) data at the bottom of the figure pose a different kind of challenge, one that dramatizes the importance of the slope-intercept methodology. Expository passages were introduced into the IRAS design relatively late in the study, and so the averages in the bottom panel are based on quite different groups of students. There is no way to tell which of the cohorts are represented at any of the aggregates plotted in the lefthand panel without resorting to a rather complex graph. In the righthand panel, by contrast, each student is represented with equal weight, and so one can compare the righthand panels for comprehension of narrative and expository prose presented orally. To be sure, the caveat still must be respected -- there are fewer data points for expository than for narrative comprehension, and so the growth rates for the latter are more trustworthy in this study.

Taking these precautions into account, it appears to us that the listening comprehension of expository passages in this study is roughly comparable to that for narrative passages. To be sure, we spent considerable energy in assuring that the two genre were equivalent in readability demands, and that both were structurally straightforward.

Finally, let us consider the three tasks that are most directly associated with "reading" -- sentence, narrative, and expository reading and comprehension. The data for these tasks are displayed in Figure 7.

The Sentence Reading subtest is shown in the upper panel. The performance measure is reading speed in syllables per second; a rate of one syllable/second or less corresponds to a slow and rather halting rate, whereas two syllables/second or faster is reasonably easy on the listener. The growth track model does not apply to this measure, neither as regards the shape (one would expect the function to eventually flatten out) or the actual growth rate. A rough projection might call for the rate to increase from a presumed level of zero at first-grade entry to a value of about 1.0 syllable at exit from first grade (the student should manage at least a halting attempt), and to have reached at least 2.0 syllables/second by third-grade entry (the student who lacks fluency in oral reading at that time will have problems with other aspects of reading, and is likely to be seen as a problem by the teacher).

The SRD data in the figure are subject to another caution. As in the case of the expository passages, the sentence reading task was added to the IRAS battery as part of the revision of IRAS following the assessment of the initial cohorts. Accordingly, the longitudinal record for this measure is not complete, and varies with the cohort. The correspondence between the year-to-year averages and the growth track projections is as a consequence not immediate, and the two graphs cannot be compared directly.

Turning now to the year-to-year averages in the figure, the general picture for the bilingual students is one of a lack of fluency on exit from first grade. The average rate of half a syllable per second (i.e., two seconds for each syllable to be pronounced) means that the typical student was experiencing considerable difficulty in oral reading of connected text. The average for the monolingual English sample is much closer to the expected performance levels exiting from first and second grade. The bilingual students do not reach comparable levels of fluency until the end of third and fourth grade.

The picture from the year-to-year averages is somewhat misleading, however, as can be seen from the slope-intercept estimates in the righthand panel. The cohorts that were tested twice only during the first and second grades are actually somewhat comparable to the monolingual English cohort in the development of oral reading skill. The contrast between these two plots illustrates the value of the slope-intercept method for drawing comparisons when there are differences between cohorts in the availability of observations.

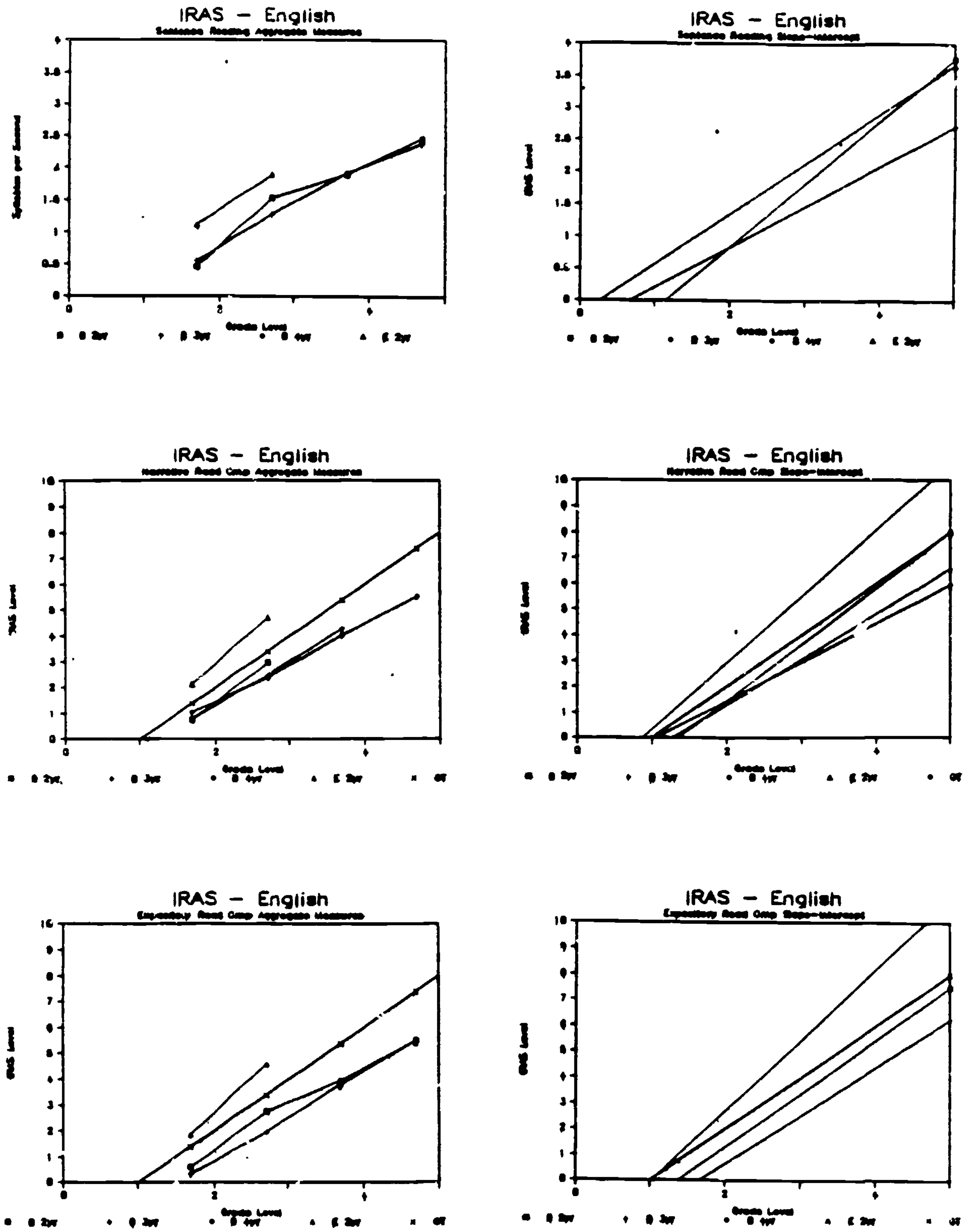


Figure 7. Growth plots of IRAS-E reading scale indices for aggregate measures (left) and best-fit regression lines (right) for individual cohorts.

The middle panel in Figure 7 shows the growth data for narrative reading comprehension. This subtest was designed according to the growth track concept, which appears on the plot as a theoretical projection. The performance of the monolingual English sample corresponds closely to the predicted values, suggesting that the design of the IRAS passages is operating as planned. The data from the bilingual samples traces out a linear path, as predicted, and the intercept estimate on entry to first grade corresponds closely to the expected value of 0.0. The growth rate, however, is less than the expectation of two IRAS levels per year. There are also differences between cohorts; the 2-year cohort actually comes close to the expected growth rate, whereas the 3- and 4-year cohorts lag behind the most. To repeat a caution made previously, the long-term projections for the 2-year cohorts are less trustworthy than those for the 3- and 4-year cohorts.

The bottom panel of the figure shows the data for expository reading comprehension. Performance on these passages conforms to the predictions of the linear growth model (recall that the expository subtests were added to IRAS after the initial testing of the first cohorts had been completed, so that the left and right panels cannot be directly compared). These passages are more complex structurally than the narrative passages, and they do appear to be somewhat more difficult for both monolingual and bilingual students to recall.

Spanish IRAS. In this section the data from the Spanish version of IRAS will be presented. The discussion is organized in a fashion exactly parallel to that used for the English version. After the Spanish-language findings have been considered, the two sets of data will be compared in the next section.

Figure 8 displays performance on the three decoding tasks. The bilingual cohorts are able to decode real words (the VDC measures in the top panel) more or less as predicted by the growth track model. The various cohorts do differ somewhat, but the degree of similarity is reasonably close. The materials used for instruction in Spanish reading are not guided quite so strictly by readability considerations, and so the fit to the growth track model is rather remarkable, both as indicated by the year-to-year averages and the slope-intercept projections.

The data for the monolingual Spanish cohort reveal performance is substantially higher than for the bilingual students. On exit from first grade the students are able to decode real words roughly corresponding to fifth grade materials. Students tested at the end of third grade are reaching the upper limits of the word lists, leading to nonlinearities and consequent insensitivity to further growth. Nonetheless, it is clear that the monolingual students are acquiring skills in the decoding of real words at a much faster rate than predicted from the reading materials used for Spanish instruction of bilingual students in the United States.

The LSC subtest for assessment of decoding skills using synthetic words had an upper limit of 4.99. This subtest could not be designed

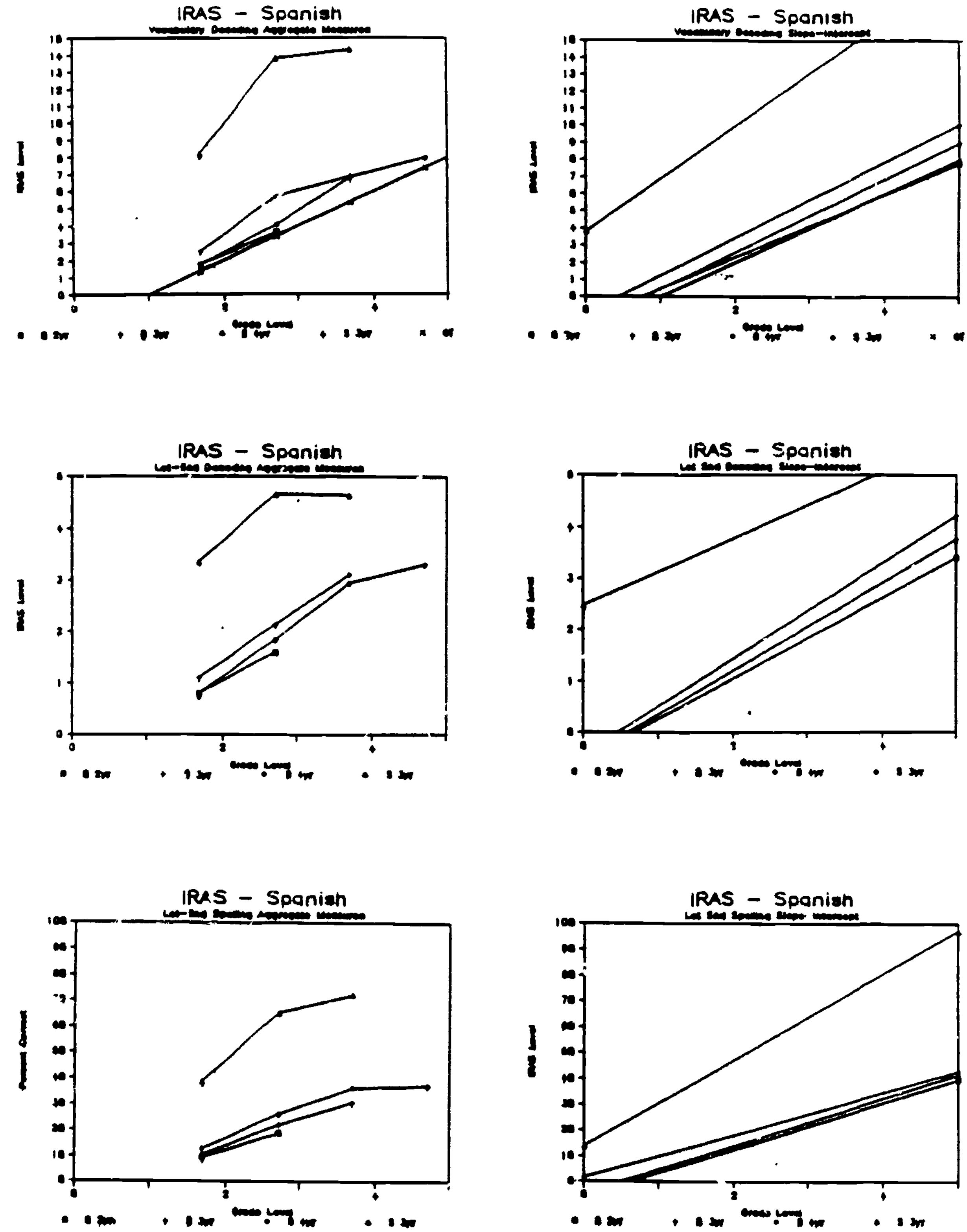


Figure 8. Growth plots of IRAS-S decoding scale indices for aggregate measures (left) and best-fit regression lines (right) for individual cohorts.

to conform to a growth track. As can be seen in the middle panel of Figure 8, however, the bilingual cohorts made steady progress over the years, reaching a level of about 3.0 at exit from fourth grade. The changes in performance are generally linear, and the differences between cohorts are negligible.

The monolingual Spanish cohort has reached a level of 3.0 by the time they leave first grade, and for practical purposes have reached the upper limit of this test by the end of second grade. Again, this finding means that the linear growth model is less appropriate for this cohort; both the slope and the intercept need to be interpreted with caution.

A different pattern shows up in the LSP subtest. As may be recalled, students were asked to spell a graded list of synthetic words, and the performance index was the proportion of words correctly spelled. A growth track cannot be projected for this measure. The data in the bottom panel of the figure show that none of the cohorts came close to perfect performance on the spelling task. Both the 4-year and the monolingual Spanish cohorts showed some nonlinear tendencies over time, but these trends were not artifactual.

Spelling of synthetic Spanish words was a difficult task for the bilingual cohorts. All groups progressed at about the same rate, attaining a level of about 40 percent correct at the end of fourth grade. The monolingual Spanish sample reached this level at exit from first grade, and then made smaller gains during the next two years, with an apparent asymptote at about 70 percent. Given their high levels of performance on the decoding subtests, it is somewhat surprising that the monolingual Spanish students encountered so much difficulty with spelling.

Figure 9 displays the data for the oral language subtests. In the top panel are the findings for the Vocabulary Definition task. The bilingual cohorts all made steady linear progress, as did the monolingual Spanish sample. VDF estimates for the bilingual students suggest that on entry to first grade they have a working vocabulary that is one or two grade levels more advanced than the text materials which they are encountering in the basal materials. All of the cohorts make slightly slower than expected progress during the primary years. The monolingual Spanish sample enters first grade with a third-grade ab-ulary, and progresses at a rate of about two IRAS levels per year -- the rate predicted by the growth track model.

The data for the listening comprehension subtests, NLC and ELC, appear in the middle and bottom panels of the figure. Recall that the year-to-year averages for ELC are based on changing groups of subjects, and so cannot be directly compared with the slope-intercept projec-tions.

All of the cohorts are estimated to enter first grade with listening comprehension skills substantially above the zero level of IRAS. The monolingual Spanish cohort performs at the third grade level

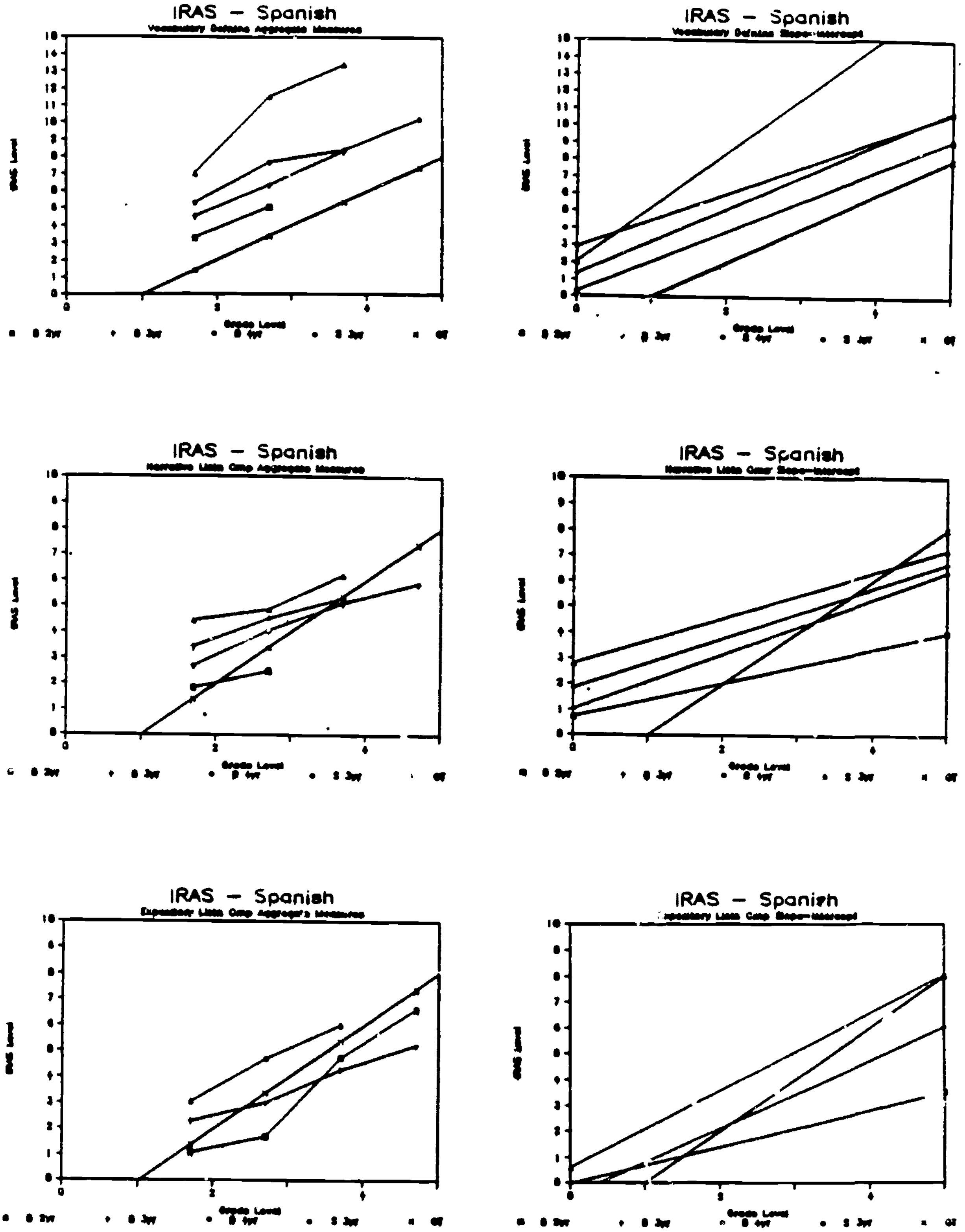


Figure 9. Growth plots of IRAS-S formal language scale indices for aggregate measures (left) and best-fit regression lines for individual cohorts.

on narrative; the bilingual cohorts are at about the second grade level. Expository comprehension is about one grade level lower than narrative. Growth appears to follow a linear course over the years, and progress is about half the rate projected by the growth track model for all cohorts. The result is that the students enter first grade able to handle passages more difficult than they will encounter in print, but by the end of third grade their listening comprehension has fallen below the projected readability levels. The comprehension growth rates for the 2-year cohort are noticeably lower than for the other groups on both subtests. Estimated entry level is higher for narrative than for expository, but the growth rates for expository are faster than for narrative; the result is that comprehension at the end of fourth grade is estimated to be roughly equal for both text genre.

Finally, Figure 10 presents the findings for the three reading subtests: sentence reading, and narrative and expository comprehension. Recall that the year-to-year findings for the SRD and ERC measures are based on different groups of students from one year to the next.

The monolingual Spanish cohort has achieved a reasonable level of fluency by the end of first grade, and progresses quite well over the remaining years. The bilingual cohorts are still at a relatively halting level on exit from second grade, and are estimated to reach a rate of two syllables per second only at the end of fourth grade. Average growth for all cohorts appears to proceed linearly, and so the linear growth model has been used to summarize performance even though the growth track model is not directly applicable.

Reading comprehension of narrative and expository passages is displayed in the middle and bottom panels of the figure. Progress is linear for all cohorts. The intercepts are generally negative -- it appears that the students are at a disadvantage in comprehending written materials on entry to first grade. In light of the results from the listening comprehension subtests, this finding cannot be due to a limitation in handling text. Growth for the monolingual Spanish sample approximates the rate projected from the growth track, but is substantially slower for the bilingual cohorts. In fact, reading comprehension for these latter students improves at less than one IRAS level per year, whereas two levels per year are expected. Expository passages appear to be somewhat more difficult than narrative texts for the youngsters in our sample.

Comparison of performance on English and Spanish IRAS. It is not our purpose in this volume to focus on substantive findings. Nonetheless, some general observations are called for. We will consider in turn the areas of decoding, formal oral language skill, and reading skill. We will look at certain relations within and between the three areas as well as between the two languages.

Decoding is an important skill to master in the primary grades. The bilingual cohorts appear on the average to be performing at a level appropriate to the demands of the texts that they are likely to

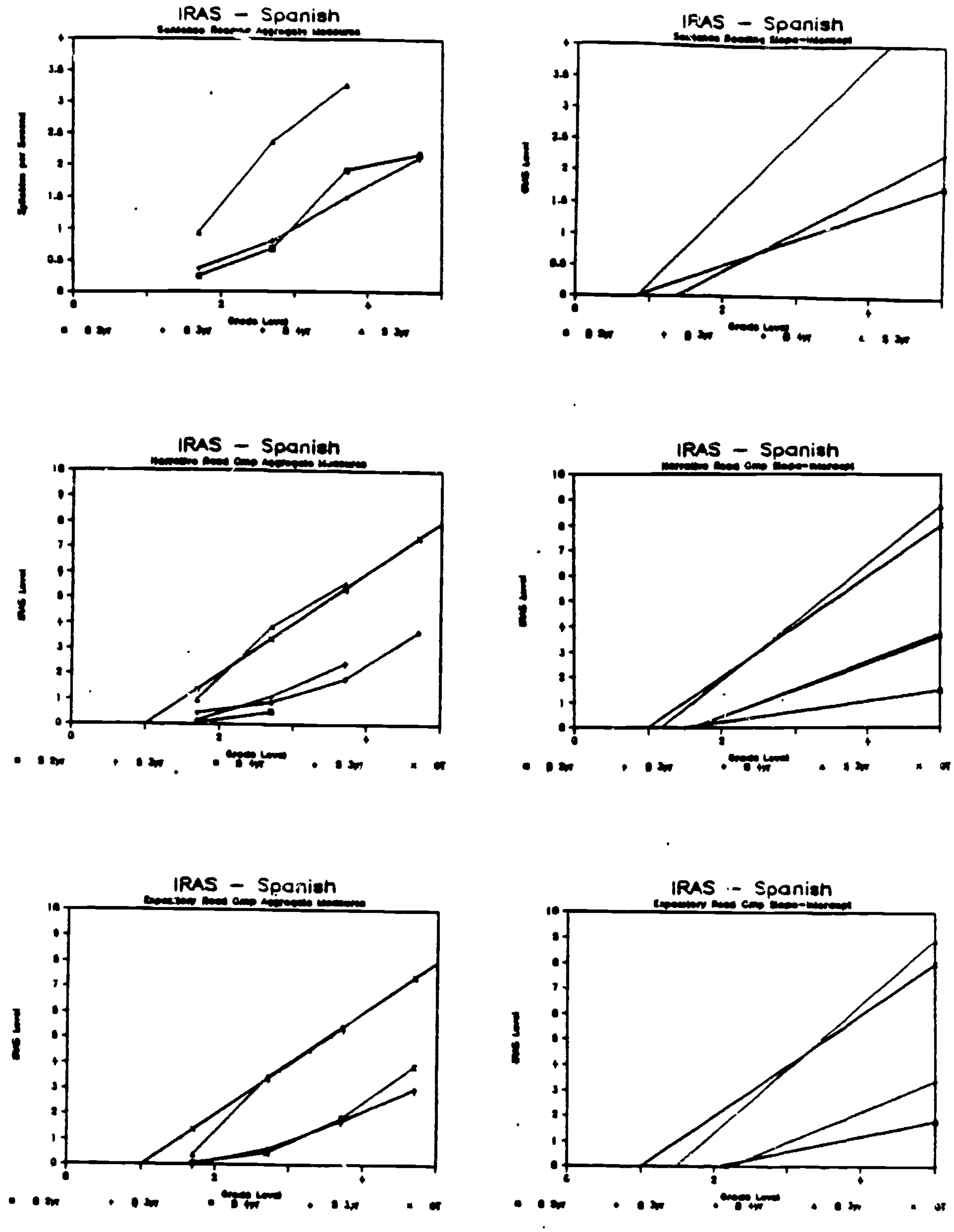


Figure 10. Growth plots of IRAS-S reading scale indices for aggregate measures (left) and best-fit regression lines (right) for individual cohorts.

encounter across the range of grades investigated in the study. The monolingual samples outperform the bilingual students, markedly so in the Mexican classrooms. We cannot estimate the level of performance for the LSC subtest, but it does appear that the bilingual students are able to transfer decoding skills to novel words that they have not encountered in their lessons. Again, they do not reach the levels of the monolingual students.

Decoding skill was assessed in the VDC and LSC subtests by asking the student to actually pronounce words (many of the standard instruments resort to a multiple-choice procedure), and so it is not proper to characterize these as "passive" in their demands on the student. Nonetheless, the spelling test entails a different kind of productivity by the student. The LSP scores for all cohorts and in both languages show that students were substantially below the upper limits of the test. Moreover, in every instance it appears that performance was reaching an asymptote beyond which no growth was observable. Even the monolingual Spanish sample, which was exceptional in the mastery of pronunciation skill, misspelled more than a quarter of the words on the average.

The bottom line, however, is that the bilingual students appear to be attaining a level of competence in decoding that should allow most of the students to handle the text demands of the typical basal series. To be sure, this conclusion entails an implicit assumption that needs to be examined -- if the student can decode a word but only with considerable attention and effort, the mental resources may not be sufficient to handle the other demands of comprehension and interpretation.

The ability of the bilingual students to handle oral language in formal situations presents a different pattern. While below the level of the monolinguals on entry to first grade, virtually all cohorts performed substantially above the level of the demands of the vocabulary and text requirements of the basal series -- in both English and Spanish! Unfortunately, this early advantage was slowly eroded over the years; the rate of growth in oral language competence was so slow for these students that by the end of fourth grade their performance was either at or below the readability of the basal materials -- the problem is much more serious for Spanish than for English.

It is also important to note that the same basic pattern holds for the monolingual cohorts. That is, they enter first grade capable of comprehending passages which they hear that are typically found in second and third grade books. The students in the English cohort gain only three IRAS levels for every four that are predicted from the growth track; the Spanish cohort progresses at about the same rate in expository listening comprehension, but drops to a rate of one IRAS level per three expected in narrative comprehension.

Several researchers have reported that comprehension is not given much systematic attention in existing reading programs. The present findings are consistent with these reports, and with the NAEP findings of poor performance on multiple-choice tests of higher-order comprehension skills at the late primary grades and in secondary school. It appears that the problem is especially serious for the bilingual students in our sample, but is also marked for the monolingual students. The difficulty cannot be directly attributed to a lack of fluent decoding. The failure of the students to progress in listening comprehension at a rate sufficient to meet text demands suggests that instruction in skilled decoding must be augmented by attention to methods for analyzing text structures.

All of the preceding strands come together in the subtests of reading skill. Oral reading performance provides an observable index of the student's ability to apply word-level decoding to connected text. The bilingual cohorts appear to be seriously lacking in this ability during the early primary grades. To be sure, the data presented in this section are averages -- it might be that some of the bilingual students are in Spanish reading programs, and attain levels of competence equal to that of the Spanish cohorts, while others in English programs match the English cohorts. Analyses in later volumes will show that this interpretation does not fit the data.

Reading comprehension subtests, where comparison with a growth track is possible, show that the bilingual students depart steadily from the expected level of performance over the primary grades in school. The gap at the end of fourth grade is about one grade equivalent in English; the average level in Spanish is at second grade or below. The monolingual cohorts perform at levels close to the projections of the growth track in virtually all instances -- to be sure, one might wonder about their ability to sustain continued growth in reading comprehension given the previous comments about listening comprehension. Nonetheless, and for reasons that will be explored in detail in subsequent volumes, the bilingual students in our sample do not appear to have been achieving a satisfactory rate of progress in reading comprehension of either narrative or expository passages.

The X-Intercept: Estimating the Onset of Instruction

In estimating the intercept of a linear function, it is usual to define an appropriate value on the X or horizontal axis as the zero or entry point. The Y-axis is passed through this point, and the best fitting line is extended to pass though the Y-axis. We will refer to the point at which the linear function strikes the Y-axis as the Y-intercept. As noted earlier, the Y-intercept estimates from the growth functions pose some difficulty in interpretation. Ideally, we would have determined for each student the point in time, either before or after entry to school, at which instruction (formal or informal) was begun on a particular component in reading, and define that point in time as the locus for estimating the intercept. Performance would presumably be zero at this locus, and on the assumption of uniform growth (presumably reflecting uniform instructional activities) the student would progress at a constant rate.

In fact, it is virtually impossible to determine the onset of instruction for individual students. One can imagine collecting data on this matter; we were not able to do this. For some components it is likely that certain students were provided instruction prior to school entry, which might pose particular problems of documentation. Our selection of entry to first grade was made on the grounds of practicality -- it may be appropriate for some components and for some students, but it is probably wide of the mark in many instances.

One consequence of this decision is that the estimates of the Y-intercept is quite often negative. We discussed earlier in this Volume some ways of thinking about a negative Y-intercept -- it might represent a level of unpreparedness by the student, perhaps due to characteristics of the preschool experiences of the child, perhaps due to the nature of the kindergarten program. We have the capability to evaluate these hypotheses to some degree, and will do so in later volumes.

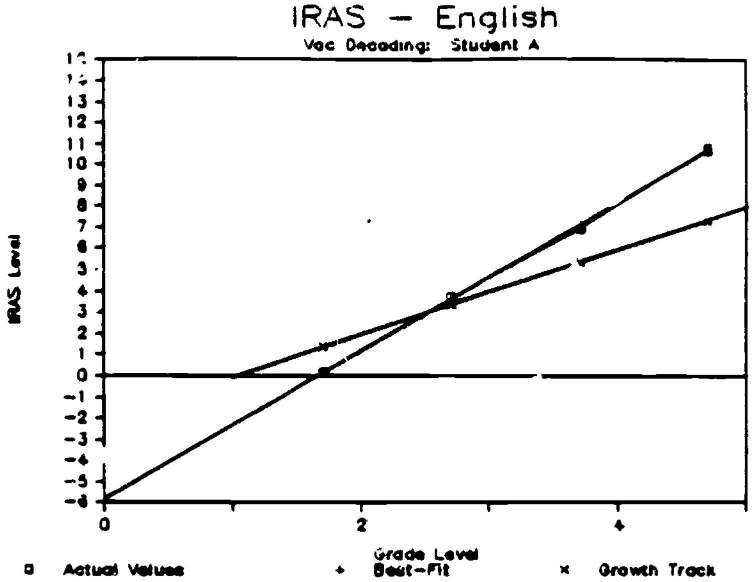
We have also explored another approach for thinking about the intercept of the linear function, however, an approach that appears sufficiently promising to merit presentation of the findings. To illustrate the concept, which we have labeled the X-intercept, let us consider again the data from the three students presented earlier in the volume for illustration of the growth-track model. These data are displayed in Figure 11 for convenience.

The Y-intercept estimates for these three students vary widely. Student C is placed almost exactly on the growth track -- a value of 0.3 at the beginning of first grade, compared with the prediction of 0.0. The estimate for student C is large and positive at 4.4, while the estimate for student A is large and negative at -2.3.

The X-intercept for each of these cases can be defined as the point at which the best-fitting linear function crosses the X-axis. The mathematics for estimating this point are straightforward, and the definition of the X-axis is not ambiguous as was true for the Y-intercept. The main question is one of interpretation; what is the meaning of the X-intercept?

We will advance the following interpretation -- the X-intercept is an estimate of the point at which instruction for a particular reading component began in an effective fashion for an individual student. Thus, for student C, it appears that training in decoding of real words began very nearly at the beginning of first grade; the X-intercept is estimated at .8. Student A, whose data might be interpreted as reflecting a relative disadvantage on entry, now can be described quite differently. Although there was a delay in the onset of instruction (the X-intercept is at 1.67, about two-thirds of the way through first grade), once effective training begins the student's response is fast and steady. For student B, the data suggest that someone (the home, or perhaps Sesame Street) began to influence the student's ability to decode familiar words; the X-intercept estimate of -1.0 indicates that these activities began at least a year prior to kindergarten entry.

BEST COPY AVAILABLE



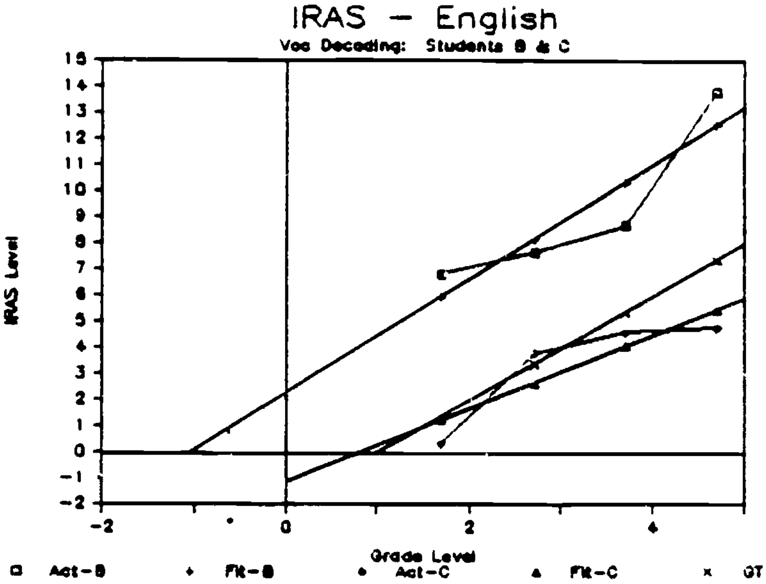


Figure 11. IRAS-E Vocabulary Decoding actual values and best-fit regression line for students A, B, and C (reproduction of Figure 3).

167

The estimation of the X-intercept is subject to the same caveats that hold for the other estimates based on the linear function. Some of the subtest indices do not lend themselves to a ready interpretation in terms of a growth track. Floor and ceiling boundaries can introduce artifactual nonlinearities that affect the various estimates. Finally, nonlinearities can arise because of inconstancies in the instructional program. Nonetheless, the use of the X-intercept as an estimate of the onset of effective instruction seems at least as informative as strict reliance on the Y-intercept estimate, and so this concept will be relied on in subsequent volumes where it seems to make sense for interpretation of student achievement.

Analysis of Deviations from the Growth Track

The growth track concept arises from an examination of the materials used for reading instruction in the United States. The findings described in the previous sections reveal a variety of departures from the projected growth track. The purpose of the present section is to consider how to account for these departures in a more or less systematic fashion. Some of the IRAS subtest indices do not lend themselves to the growth track concept. Other measures to be discussed later in this volume will be subjected to analysis by the linear change model, though not necessarily through the same argument used for IRAS. The present discussion holds for all of these analyses with appropriate modifications.

Aggregate Deviations

If one examines the year-to-year averages present above, it is clear that the projected growth track gives an accurate account of the aggregate values for some of the IRAS components, but is wide of the mark for others. In accounting for observed performance, the first task might be to provide a reason for these departures. In fact, the closing remarks of the previous section comprise just such an effort. By way of a brief reprise: (a) the readability or word frequency formulas that constitute the major design considerations constraining the development of textual materials in basal series, and which are the primary basis for the growth tracks, are closely captured in the vocabulary decoding tasks; (b) where the student's familiarity with oral language is tapped by an IRAS subtest, it is obvious that the readability constraints, which are applied across the board (decoding, word meaning, and text comprehension), mean that the student is presented with materials substantially below the student's functional capability; and (c) in a few instances, departures in the year-to-year averages arise because students have reached the upper limits of a particular subtest. In any event, the analysis of departures from the projected growth track for a given subtest should probably begin with an examination of the differences between the predicted level and the overall average.

Deviation of Individual Students from the Aggregate

Let us next consider a subtest index in which aggregate performance is reasonably linear (this state of affairs holds for quite a few of the IRAS measures), and where the aggregate either corresponds to the projected growth track or we have a reasonable explanation for the overall departure (again, not an unreasonable precondition). We will still find a considerable amount of variation in the slope and intercept estimates for individual students around the aggregate values; either the overall average or the averages of the slope and intercepts may be taken as the frame of reference for the deviatic of individual students. How are these departures from a standard to b. accounted for?

This question, which focuses on individual differences between students, is a time-honored problem for educational psychology. The present study places the matter in a somewhat different perspective on occasion, but the general plan of attack seems clear enough. Let us begin by separating the entry level estimates (the Y-intercepts) from the growth rate estimates (the slopes).

Why are there differences between students in their competence in some area of academic skill at the time that they enter school? The home, the community, personal and demographic characteristics -- each of these sources of predictive factors have been extensively employed to provide an answer to the previous question.

Why are there differences between students in the rate at which they learn once they enter school? Interestingly, the same set of factors is used to explain such deviations on those occasions when the question is framed as clearly as it is in this study. Students can benefit from the resources provided by the school to the extent that the home gives support to the efforts of the school. In addition, the school, the teacher, and the instructional program may also influence the student's rate of progress. This latter source of influence has been subject to considerabl debate. Nonetheless, it does seem that an argument can be mounted that some schools are more effective than others, that there are teaching practices that promote higher rates of academic progress, and that some approaches to reading instruction lead to greater improvement (on relevant measures of reading) than do others.

These comments have implications for the stage of analysis subsequent to consideration of aggregate differences from the projected growth track. For each of the IRAS indices (and for other measures with similar characteristics), both the growth rate and the intercept estimates, it is possible to list a set of predictive factors appropriate to the particular index. Standard methods of regression analysis can then be used to render an account of the observed deviations of students from th average of the index.

It should be noted that the design of our assessment package and the way in which we have summarized performance does have implications for the selection of predictive factors. By decomposing reading into a set of relatively precise components, we have made it easier to identify instructional elements corresponding to student performance. When reading is measured in an omnibus fashion, it is not easy to specify what to look for in classroom practice or in the basal materials that might promote growth in the component. The design of IRAS makes such identification easier, and so inclines us to rely more on predictors from the classroom than those that reflect the background characteristics of the student. In like fashion, if we were to employ the Y-intercept exclusively in the analysis (standard practice, as it were), again the background characteristics of the students would take on primary importance. By including the X-intercept estimates in the analysis, it becomes more natural to look at the data from classroom practices to investigate the possibility that instruction may have been delayed for some students relative to others.

Deviations from Linearity

The preceding discussion rests on the more or less implicit assumption that average performance over the grades is linear, or the achievement of the individual student proceeds in a constant (i.e., linear) fashion, or both. In this section both of these assumptions will be examined more carefully.

The first matter can be dispensed with rather quickly. Most of the IRAS indices were designed to reflect student progress in a linear fashion, assuming that the design of the typical basal series was a trustworthy guide for learning. While the latter assumption was not always realized, in fact most of the IRAS measures do change in a reasonably straightline fashion on the average. The few exceptions were either not expected to fit the linear model, or were influenced by ceiling artifacts. These instances, few in number, are not the primary concern of this section, and will be handled as individual cases at the appropriate time.

More interesting are situations like those for students B and C who served as illustrations previously in the volume. Performance on the Vocabulary Decoding task is linear on the average, and for many individual students the growth track is remarkably close to a straight line. How are we to account for the apparent nonlinearities in protocols like those for students B and C?

There is always the possibility that these departures are simply reflective of the unreliability that is always present in achievement tests. To the degree that this argument holds, then we will have no success in predicting deviations from linearity.

Another possibility is to be found in developmental changes in the student -- student B may have suddenly advanced in vocabulary decoding in fourth grade because of a shift to a level of formal operational thought, if we may take a Piagetian stance for the moment. Our data do not permit a test of this hypothesis, and so again we should have no success in predicting deviations from strict linearity. More generally, one can imagine a number of contextual factors that might affect student performance from one test point to another -- a bad day, a divorce in the family, an argument with the teacher -- that would lead to a departure from what would otherwise be a constant rate of progress. All such deviations fall into the same category -- unless there are data that speak to these conditions, then the deviations will be unpredictable.

Our primary hypothesis about deviations from linearity is much less subtle, and rests on data sources that are available to us. We propose that the student's rate of progress during a given year may depend on the program of reading instruction in effect during that year. If the teacher emphasizes reading, if the concentration is on decoding or comprehension, if textual materials are obviously present, if time is well used, if the classroom is orderly and well managed, then we would predict a positive "bump" in student growth -- an upward deviation from the best-fitting linear growth function. If one finds the reverse of these conditions, then a negative deviation would be expected.

This hypothesis leaves many questions unanswered. How shall we combine general facets of classroom practice with the curriculum specifics? What if the classroom is well managed, but comprehension is neglected? How should we take into account the student's profile on entry to a particular grade? A given program of instruction may be just what one student needs, but may fail another student for a variety of reasons.

All such concerns are relevant, and to the degree that they are important but unanswered by our method of analysis, then once again we will be unable to predict individual departures from strictly linear growth. Nonetheless, since this approach seems most promising and sensible to us, it will provide the basis for our efforts to account for departures from constant growth for individual students.

References

Calfee, R. C. (1982). Literacy and illiteracy: Teaching the nonreader to survive in the modern world. Annals of Dyslexia, 32, 71-91.

Calfee, R. C., & Calfee, K. H. (1981). Interactive reading assessment system (IRAS). Unpublished manuscript (rev.). Stanford, CA: Stanford University.

Calfee, R. C., & Curley, R. G. (1984). Structures of prose in the content areas. In J. Flood (Ed.), Understanding reading comprehension. Newark, DE: International Reading Association.

Calfee, R. C., Calfee, K. H., & Pena, S. C. (1979). Interactive reading assessment system-Spanish. Unpublished manuscript. Stanford, CA: Stanford University.

Cronbach, L. J., & Furby, L. (1970). How should we measure "change"--Or should we? Psychological Bulletin, 74(1), 68-80.

Estes, W. K. (1956). The problem of inference from curves based on group data. Psychological Bulletin, 53, 134-140.

Greeno, J. G. (1980). Psychology of learning, 1960-1980: One participant's observations. American Psychologist, 35, 713-728.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. Psychological Review, 85(5), 363-394.

Kintsch, W., & Yarbrough, J. C. (1982). Role of rhetorical structure in text comprehension. Journal of Educational Psychology, 74, 828-834.

Klare, G. R. (1974-75). Assessing readability. Reading Research Quarterly, 10, 62-102.

Rogosa, D. R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. Psychological Bulletin.

Rogosa, D. R., & Willett, J. B. (In press). Understanding correlates of change by modeling individual differences in growth. Psychometrika.

Venezky, R. L. (1970). The structure of English orthography. The Hague: Mouton.

APPENDIX A

Descriptive Statistics for Each of the IRAS-E and IRAS-S Scales for Individual Cohorts

Table 1

Interactive Reading Assessment System - English:
Descriptive Statistics for Decoding Scales
for Individual Cohorts

Scale	Bilingual	Statistic	Year 1	Year 2	Year 3	Year 4	Y-Intl	Slope
VDC	2-years	M	2.01	5.4[illegible]			-0.39	3.43
		S	2.40	3.25			2.92	2.27
		N	152	152			152	152
	3-years	M	1.92	5.42	7.78		0.06	2.91
		S	2.20	4.02	4.29		2.42	1.61
		N	38	37	37		38	38
	4-years	M	2.94	4.77	8.07	10.1[illegible]	1.02	2.48
		S	2.47	3.50	4.47	4.12	2.75	1.03
		N	56	56	56	56	56	56
	Mono-Eng							
	2-years	M	4.63	6.94			3.02	2.31
		S	2.80	2.70			3.98	2.53
		N	36	36			36	36
LDC	2-years	M	1.57	3.63			0.12	2.06
		S	2.34	2.36			3.34	2.20
		N	149	149			149	149
	3-years	M	1.14	3.22	4.66		-0.0[illegible]	1.78
		S	1.82	2.68	2.24		2.34	1.07
		N	38	37	37		38	38
	4-years	M	1.[illegible]1	3.04	4.75	5.17	0.46	1.39
		S	1.51	2.39	2.37	2.18	1.87	0.72
		N	56	56	56	56	56	56
	Mono-Eng							
	2-years	M	3.36	4.95			2.25	1.59
		S	2.54	1.83			3.61	1.93
		N	36	36			36	36
LSP	2-years	M	9.30	24.22			-0.81	14.72
		S	14.36	19.91			18.05	14.57
		N	146	146			146	146
	3-years	M	8.13	22.28	29.52		2.90	9.88
		S	15.10	23.06	21.52		17.56	7.52
		N	41	40	39		42	42
	4-years	M	9.87	19.49	33.38	40.90	2.37	10.70
		S	12.60	17.34	24.40	22.77	15.60	6.57
		N	52	52	52	52	52	52
	Mono-Eng							
	2-years	M	27.84	35.99			22.13	8.15
		S	18.82	21.36			23.92	15.60
		N	36	36			36	36

Note: Tabled values are based on critical indices where one IRAS unit equals .5 grade-levels (except for LDC and LSP).

Table 2

Interactive Reading Assessment System - English:
Descriptive Statistics for Formal Language Scales
for Individual Cohorts

Scale	Bilingual	Statistic	Year 1	Year 2	Year 3	Year 4	Y-Intl	Slope
VDF	2-years	M	5.02	6.95			3.67	1.93
		S	2.38	1.86			3.83	2.49
		N	149	149			149	149
	3-years	M	4.53	5.96	7.52		3.34	1.61
		S	3.13	3.02	3.30		3.57	1.56
		N	39	39	37		39	39
	4-years	M	4.83	6.15	8.01	9.88	3.47	1.70
		S	3.17	3.18	3.12	3.20	3.52	1.00
		N	55	55	55	55	55	55
	Mono-Eng 2-years	M	7.06	8.12			6.32	1.05
		S	1.63	1.93			2.62	2.04
		N	36	36			36	36
NLC	2-years	M	3.51	5.27			2.29	1.76
		S	2.15	1.55			3.17	1.80
		N	152	152			152	152
	3-years	M	3.00	4.49	5.74		2.09	1.36
		S	2.25	2.38	1.87		2.79	0.95
		N	38	37	37		38	38
	4-years	M	2.94	4.16	5.60	6.50	1.31	1.48
		S	1.53	2.29	2.05	1.54	2.01	0.51
		N	56	56	56	56	56	56
	Mono-Eng 2-years	M	4.86	6.26			3.89	1.39
		S	1.60	1.26			2.50	1.60
		N	36	36			36	36
ELC	2-years	M	2.59	4.66	5.49	6.55	1.37	1.95
		S	2.21	2.10	2.39	1.96	3.36	1.93
		N	149	153	23	21	173	173
	3-years	M	2.35	3.37	5.34	6.38	0.56	1.71
		S	2.06	2.39	2.06	1.71	3.29	0.94
		N	33	70	70	37	70	70
	4-years	M						
		S						
		N						
	Mono-Eng 2-years	M	4.08	6.29			2.53	2.21
		S	2.03	1.26			3.00	1.62
		N	36	36			36	36

175

BEST COPY AVAILABLE

Table 3

Interactive Reading Assessment System - English:
Descriptive Statistics for Reading Scales
for Individual Cohorts

Scale	Bilingual	Statistic	Year 1	Year 2	Year 3	Year 4	Y-Intl	Slope
SRD	2-years	M	0.47	1.53	1.91	2.46	-0.15	0.98
		S	0.67	0.93	1.10	1.14	1.04	0.72
		N	149	153	23	21	173	173
	3-years	M	0.56	1.28	1.92	2.39	0.21	0.63
		S	0.60	0.97	1.01	1.11	1.01	0.34
		N	33	70	70	37	70	70
	4-years	M						
		S						
		N						
	Mono-Eng 2-years	M	1.11	1.89			0.57	0.78
		S	0.79	0.58			1.11	0.60
		N	36	36			36	36
NRC	2-years	M	0.77	2.97			-0.77	2.20
		S	1.47	2.20			2.10	1.87
		N	152	152			152	152
	3-years	M	0.83	2.50	4.28		-0.47	1.77
		S	1.31	2.44	2.50		1.67	0.96
		N	38	37	37		38	38
	4-years	M	1.05	2.39	4.04	5.34	-0.07	1.51
		S	1.29	2.16	2.77	2.34	1.61	0.66
		N	56	56	56	56	56	56
	Mono-Eng 2-years	M	2.14	4.72			0.33	2.58
		S	2.05	2.21			2.63	1.64
		N	36	36			36	36
ERC	2-years	M	0.62	2.76	3.96	5.31	-0.78	2.07
		S	1.33	2.29	3.26	2.87	2.94	1.87
		N	149	153	23	21	173	173
	3-years	M	0.33	1.99	3.83	5.32	-1.23	1.87
		S	0.89	2.34	2.78	2.47	2.88	1.13
		N	33	70	70	37	70	70
	4-years	M						
		S						
		N						
	Mono-Eng 2-years	M	1.87	4.61			-0.05	2.74
		S	1.94	2.36			2.40	1.66
		N	36	36			36	36

Note: Tabled values are based on critical indices where one IRAS unit equals .5 grade-levels (except for SRD).

Table 4

Interactive Reading Assessment System - Spanish:
Descriptive Statistics for Decoding Scales
for Individual Cohorts

Scale	Bilingual	Statistic	Year 1	Year 2	Year 3	Year 4	Y-Intl	Slope
VDC	2-years	M	1.77	3.61			0.48	1.84
		S	3.39	4.69			4.15	3.32
		N	152	152			152	152
	3-years	M	2.52	5.69	6.88		1.23	2.20
		S	4.37	6.38	5.67		5.03	2.10
		N	40	39	39		40	40
	4-years	M	1.78	4.05	6.83	7.95	0.46	2.13
		S	3.11	4.91	5.44	5.37	3.61	1.52
		N	56	56	56	56	56	56
	Mono-Span							
	3-years	M	8.19	13.84	14.36		6.69	3.09
		S	5.29	2.48	2.19		6.24	2.50
		N	37	37	37		37	37
LDC	2-years	M	0.82	1.61			0.27	0.79
		S	1.61	1.99			2.06	1.47
		N	147	147			147	147
	3-years	M	1.11	2.15	3.12		0.51	0.93
		S	1.76	2.16	1.83		2.22	0.91
		N	41	41	40		42	42
	4-years	M	0.82	1.85	2.95	3.30	0.33	0.85
		S	1.59	2.04	1.94	1.71	1.94	0.63
		N	54	54	54	54	54	54
	Mono-Span							
	3-years	M	3.35	4.66	4.64		3.12	0.65
		S	1.92	0.39	0.56		2.21	0.90
		N	37	37	37		37	37
LSP	2-years	M	9.74	18.80			3.51	9.07
		S	18.38	24.13			21.70	16.19
		N	149	148			149	149
	3-years	M	10.59	22.12	30.78		4.79	9.23
		S	17.29	24.00	25.17		18.73	9.03
		N	43	42	40		43	43
	4-years	M	13.03	26.24	36.54	37.09	10.08	8.23
		S	21.79	25.43	23.64	25.66	23.59	[illegible]
		N	52	52	52	52	52	52
	Mono-Span							
	3-years	M	59.01	65.43	72.47		30.53	16.73
		S	26.69	15.84	16.35		29.59	10.27
		N	36	36	36		36	36

Note. Tabled values are based on critical indices where one IRAS unit equals .5 grade-levels (except for LDC and LSP).

177

Table 5

Interactive Reading Assessment System - Spanish: Descriptive Statistics for Formal Language Scales for Individual Cohorts

Scale	Bilingual	Statistic	Year 1	Year 2	Year 3	Year 4	Y-Intl	Slope
VDF	2-years	M	3.29	5.07			2.06	1.77
		S	3.29	3.74			4.46	3.09
		N	149	151			151	151
	3-years	M	4.53	6.33	8.33		3.24	1.91
		S	3.73	3.60	3.01		4.47	1.92
		N	39	53	53		54	54
	4-years	M	5.30	7.66	8.43	10.23	4.48	1.56
		S	3.46	4.08	3.68	3.19	3.71	0.83
		N	40	40	40	40	40	40
	Mono-Span 3-years	M	7.02	11.54	13.42		5.22	3.20
		S	4.67	2.64	0.76		5.89	2.28
		N	13	13	13		13	13
MLC	2-years	M	1.84	2.50			1.38	0.66
		S	1.95	2.21			2.37	1.47
		N	152	152			152	152
	3-years	M	3.44	4.53	5.35		2.81	0.97
		S	2.27	2.09	1.54		2.69	1.02
		N	40	39	39		40	40
	4-years	M	2.68	4.06	5.15	5.89	2.09	1.07
		S	1.28	1.65	1.75	1.47	1.64	0.54
		N	56	56	56	56	56	56
	Mono-Span 3-years	M	4.45	4.88	6.21		3.68	0.89
		S	1.62	2.03	2.01		[illegible]	1.93
		N	37	37	37		[illegible]	37
SLC	2-years	M	1.10	1.72	4.76	6.64	0.68	0.71
		S	1.62	2.09	1.59	1.20	2.61	1.58
		N	149	155	23	21	173	173
	3-years	M	2.32	3.01	4.30	5.23	0.79	1.31
		S	1.94	2.31	2.06	1.66	3.13	1.06
		N	35	72	72	37	72	72
	4-years	M						
		S						
		N						
	Mono-Span 3-years	M	3.05	4.73	6.04		2.08	1.49
		S	2.04	2.24	1.97		2.42	0.98
		N	37	37	37		37	37

Table 6 BEST COPY AVAILABLE

Interactive Reading Assessment System - Spanish:
Descriptive Statistics for Reading Scales
for Individual Cohorts

Scale	Bilingual	Statistic	Year 1	Year 2	Year 3	Year 4	Y-Intl	Slope
SRD	2-years	M	0.25	0.69	1.93	2.19	0.09	0.42
		S	0.43	0.89	1.40	1.43	1.24	0.63
		N	149	153	23	21	173	173
	3-years	M	0.37	0.81	1.50	2.13	-0.23	0.63
		S	0.66	0.92	1.06	1.07	1.03	0.44
		N	35	72	72	37	72	72
	4-years	M						
		S						
		N						
	Mono-Span 3-years	M	0.93	2.33	3.29		0.20	1.18
		S	0.63	0.77	0.80		0.76	0.37
		N	37	37	37		37	37
NRC	2-years	M	0.10	0.32			-0.20	0.42
		S	0.46	1.21			[illegible]	1.06
		N	152	152			152	152
	3-years	M	0.18	1.12	2.43		-0.65	1.10
		S	0.56	1.88	2.39		1.00	1.08
		N	40	39	39		40	40
	4-years	M	[illegible]	0.98	1.82	3.69	-0.62	1.06
		S	0.98	1.47	2.25	[illegible]	1.[illegible]	0.78
		N	56	56	56	56	56	56
	Mono-Span 3-years	M	0.99	3.88	5.39		-0.43	2.30
		S	1.24	2.20	2.20		1.75	1.13
		N	37	37	37		37	37
ERC	2-years	M	0.07	0.49	1.85	3.89	-0.61	0.61
		S	0.41	1.15	2.33	3.68	2.51	1.38
		N	149	153	23	21	173	173
	3-years	M	0.00	0.64	1.74	2.99	-1.46	1.21
		S	0.00	1.52	2.19	2.38	1.69	1.10
		N	35	72	72	37	72	72
	4-years	M						
		S						
		N						
	Mono-Span 3-years	M	0.39	3.48	5.46		-1.20	2.53
		S	0.99	2.54	2.23		1.32	1.10
		N	37	37	37		37	37

Note: Tabled values are based on critical indices where one IRAS unit equals .5 grade-levels (except for SRD).

1

Final Report

TEACHING READING TO BILINGUAL CHILDREN STUDY

Volume 4
Oral Language Growth

Betty J. Mace-Matluck
Wesley A. Hoover
Robert C. Calfee

Document BRS-84-R. 1-IV

Preston C. Kronkosky
Executive Director

Southwest Educational Development Laboratory
211 East Seventh Street
Austin, Texas
(512) 476-6861

November 1984

Final Report

TEACHING READING TO BILINGUAL CHILDREN STUDY

Volume 4
Oral Language Growth

Betty J. Mace-Matluck, Wesley A. Hoover
Robert C. Calfee

Document BRS-84-R. 1-IV

Preston C. Kronkosky, Executive Director
Southwest Educational Development Laboratory
Austin, Texas

November 1984

There were many individuals and institutions who contributed to this research effort. We wish to express our sincere gratitude to the parents, students, and school personnel who provided the necessary data from which this study is derived.

In addition, several other individuals made valuable contributions to the study, for which we are indebted: Robert C. Calfee, Sylvia C. Peña, and Blanca de Alvarez.

And finally, we wish to thank the local data collectors at the school sites, many of whom remained with the study throughout its duration: Ramiro Barrera, Beatrice Cantú. Irene Cavazos, Carolyn Cruz, María de Obregón, Gloria de Torres, Gig Galván, Olga Hernández, Irene Méndez, Guadalupe Treviño, Rosalinda Villanpando, and Gloria Villarreal. Their patience, dedication, and hard work helped make this study a reality.

Betty J. Mace-Matluck
Wesley A. Hoover

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure | Page

185

PREFACE

In June 1978 the National Institute of Education (NIE) funded the Southwest Educational Development Laboratory (SEDL) to conduct a longitudinal study on the Teaching of Reading to Bilingual Children. Educators and policymakers alike have long recognized that the ability to read is essential for success in school, in work, and in life; yet many children from second-language backgrounds have trouble learning to read in schools today. The majority of these youngsters are from Spanish-language backgrounds and from low income families. Special programs designed to meet the needs of these children are provided in schools, but there is limited research evidence to guide the development, evaluation, and implementation of these programs. This study is intended to provide information that will result in greater insights into what constitutes a favorable learning environment for children from Spanish-language backgrounds, what instructional sequences and events promote successful and efficient learning of literacy skills, and what the language and literacy outcomes of current schooling practices are for a large sample of these youngsters.

The study was conducted during the years of 1978 through 1984. It is a comprehensive longitudinal investigation of the development of reading skills from kindergarten through fourth grade for a representative sample of more than 350 children from bilingual backgrounds, and for smaller samples of children who, on entry into school, were monolingual in English or Spanish. In this "natural variation" study, teaching and learning were carefully documented in field settings at the several sites.

The goals of the study were to (a) describe variations in both English and Spanish language ability of students living in bilingual communities, (b) document prevailing practices in reading instruction for bilingual students, and c) investigate the relations between the instructional program and student achievement for students with differing entry profiles.

Description of the Study

Surveys of the general and school populations reveal an increase in the number of students whose language resources are not an ideal match to the language of the school. An important question for educational practice and policy centers around the school's responsibilities in this situation. Bilingual programs, English-as-a-Second-Language classes, classroom aides, and "sink-or-swim" approaches can all be found in practice today. From limited evidence now available, none of these techniques has emerged as the one best system.

Hispanics make up the largest and fastest growing school-age population today. The demographics for some states show that over the next decade they may constitute as much as a third to a half of the population. In the state of Texas at present approximately one third of the school children are from Hispanic backgrounds (approaching one

illion). They are found in virtually ever school district in the state. Many of the school districts in the southern portion of the state serve school populations of which 75% to 99% of the children are from Spanish-speaking backgrounds and, on entry into school, are often limited in their ability to speak English and to profit from instruction in that language. This population is not restricted to the border areas, however. Large urban centers in the state report as much as 20% of their school population from Hispanic backgrounds, with a concentration of some 80% to 90% in certain of their schools.

It is well documented that, in general, children from Spanish-speaking backgrounds, for whatever reason, often encounter difficulty in our nation's schools: they do more poorly on standardized tests than does the general school population, and their dropout rate is high. Bilingual education, in which students are given instruction partially through the home language until they have attained sufficient proficiency in English to benefit from English-medium instruction, has been the principal approach recommended by the Office for Civil Rights to ensure access to equal educational opportunity for these children. Although many individual programs have had considerable success in improving the academic performance of language-minority students, it has not been demonstrated that these programs generally are reducing inequality of educational opportunity on the large scale that was envisioned.

Growth in reading comes about for most youngsters through formal classroom instruction. Understanding the development of reading, and knowledge of the critical variables that determine success or failure, depends on a careful examination of the instructional program -- not just the label over the classroom door, but the program as actually implemented by the classroom teacher.

Educators have raised several issues about the most effective way to help bilingual children become proficient readers of English. These include (a) valid assessment of the student's ability in the languages of the home and of the school, (b) the optimal balance of formal instruction in both languages, (c) the most effective trans[illegible] from one language to the other, and (d) bilingual support within [illegible] classroom environment. A major thesis of the Teaching Reading to Bilingual Children study is that addressing these issues (and others) requires a comprehensive and ecologically-valid investigation of the linkage between the child's language and the language of instruction.

Design of the Study

To achieve the objectives of the study, considerable attention was given to the selection of schools, teachers and students, to the instruments for assessing language and reading achievement, and to the methods for evaluating the classroom instruction. Each of these topics is discussed briefly below.

Schools, Classes and Teachers

Twenty schools and 200 teachers from six school districts participated in the study. Included are variations in the nature of the reading program (a range from phonics-oriented to meaning-based), classroom organization (some self-contained, others team-taught), and grade structure (the range of grades in the individual school and the extent of cross-grading both vary). The schools differed in size, SES, urbanicity, locale, and makeup of the student body (from medium to high concentration of bilingual students).

Student Cohorts

The study was undertaken in four cohorts or "waves" of students. Three of the cohorts consisted entirely, or in large part, of bilingual students. The first cohort was small (N=40) and of limited generality; the second was somewhat larger (N=80) and covered a slightly broader array of contexts. The third cohort which was both larger (N=200) and broader in its generality, incorporated a number of procedural improvements based on previous experience in the study and included a monolingual English-speaking sample. The fourth cohort consisted of a relatively small sample (N=60) of monolingual Spanish-speaking students.

All of the bilingual sites were from the state of Texas, as were the monolingual English-speaking students. The monolingual Spanish-speaking students were from one site in Northern Mexico.

The original design of the study called for each student to be assessed and observed from entry to kindergarten through exit from third grade. By covering the full range of the primary years, we would be able to examine the transition from "learning to read" through "reading to learn." For students in programs where the initial stages of reading were in Spanish, we also considered it important to determine the transition to competence in English reading.

The original design was in fact implemented for the first two cohorts; some of the students were tracked from first through fourth grade, but most followed the intended design. Due to limited funding in the later stages of the study the last two cohorts could not be followed for the full four years that were originally intended. The bilingual and monolingual English samples from the Texas sites were observed from kindergarten through second grade, and the monolingual Spanish samples from the site in Northern Mexico were observed from first through third grade (the program did not provide a kindergarten).

The monolingual samples were incorporated in the design to aid in validating the instruments for student assessment. Both the English and Spanish cohorts are small and not selected to be fully representative of monolingual populations. Data from these samples will be presented in Volume 3, as part of the discussion on the adequacy of the instruments for measuring growth. The study was designed to study the course of reading in bilingual students, not as a basis for comparing these students with monolingual youngsters. Accordingly, comparisons

between the various samples will not be made in this report, nor do we recommend that others attempt such comparisons.

Language Assessment

Several types of data were collected for each student on English and Spanish proficiency. Each year, early in the Fall and again in the Winter and Spring, teachers rated their students' language skills. Oral language proficiency tests were administered in the Fall of each year. Finally, audiotaped speech samples were obtained monthly on a rotating schedule in three settings: in the classroom, on the playground, and in the home.

Reading Assessment

Several instruments were used to measure reading achievement. Standardized test scores (mostly English) were collected yearly. More detailed information was obtained from a battery of individually-administered "performance based tests" in both English and Spanish. In kindergarten, the Stanford Foundation Skills Test was employed to measure the child's pre-reading skills. From the end of first grade on, the Interactive Reading Assessment System was administered during the Spring of each school year. This instrument provides independent measures of the student's skills in decoding, word meaning, fluency in oral reading, and comprehension. Finally, informal reading inventories were administered throughout the school year.

Classroom Observations and Teacher Interviews

Project staff conducted monthly observations of the reading instruction in each classroom and interviewed the teachers quarterly about their instructional plans. The observation instrument documented staffing patterns, grouping and organization, time allocation, the language of instruction, the character of instruction, the materials and procedures used, and the response of the students. The interviews focused on the teacher's general instructional objectives, as well as the objectives for individual target students. Taken together, these two instruments yield a rich characterization of the classroom environment for the target students.

Student Entry Variables, Classroom Factors, and Reading Achievement

The primary goals of the analyses were to identify the general relationships that characterize variation in these factors and to look for underlying regularities that are associated with success and failure, both in the early stage of reading instruction and in the year-to-year variations.

Documents

This report is one of a series of eight documents contained in the Final Report submitted to the National Institute of Education. A com-

plete list of these documents is provided on the inside of the cover of this report.

The study was a collaborative effort among a number of individuals and institutions. All members of the research team contributed to the thinking, planning, and writing of this series of documents, however, the individual whose name appears first in the list of authors was responsible for preparing the particular document.

Betty J. Mace-Matluck
Wesley A. Hoover
Co-Principal Investigators

Austin, Texas
November 30, 1984

Introduction

For the purpose of assessing the students' language abilities and monitoring their language growth, three types of language measures were employed in the study: (a) oral language proficiency tests, (b) teacher ratings, and (c) audiotaped interactions - language samples.

Each of the measures is discussed below, providing details of the tasks, materials, scoring, reliability, and descriptive statistics on the sample's performance.

Oral Language Proficiency Measures

Oral language proficiency tests have been widely used in school districts in the state of Texas since 1973, at which time bilingual education or special language programs were mandated by state law (Senate Bill 121). Home-language surveys and scores on oral language proficiency tests have been the principal means by which students are identified for special language assistance programs. State policy requires that oral language proficiency tests used in Texas schools be selected by the school district from a list of state-approved, commercially-available language tests (Texas Education Agency, 1978; 1981). In four of the five Texas sites included in the study, the Language Assessment Scales - LAS (De Avila & Duncan, 1977) was administered in both English and Spanish in the Fall of each year to students who, on initial entry into the district, were identified as potential limited English-speaking students by the Home Language Survey. The English version of the LAS was readministered in subsequent years to students enrolled in bilingual or special language programs for the purpose of determining readiness for transfer to the mainstream programs. The remaining Texas district (Site 5) administered the Bilingual Syntax Measure - BSM (Burt, Dulay & Hernández-Chávez, 1973) at the kindergarten level, and used the LAS at subsequent grade levels. In the Northern Mexico site, the Batería Woodcock de proficiencia en el idioma - Version en Español (Woodcock, 1981) was administered to the students in the study during their second and third grade years.

Data reported in this document focus on the language growth and development of the bilingual sample in the Texas schools. As noted above, the principal oral language proficiency test used by the Texas schools was the LAS; a detailed description of this instrument is presented below. For a description of the other two instruments used by schools in the study, see Appendix A.

Language Assessment Scales

The following sections provide a detailed description of the English and Spanish versions of the LAS, the scoring procedures followed in obtaini summary measures of performance, the results of reliability assessments of the two language versions, and descriptive statistics on the growth patterns shown by the target sample.

Tasks, Materials, and Scoring

The Language Assessment Scales - LAS was designed "to assist school personnel to identify children with oral English language difficulties and to ascertain linguistic proficiency in English and Spanish at elementary and secondary levels" (Duncan & De Avila, 1981, p. 1). There are two levels of both the English and Spanish versions. Students in the SEDL Bilingual Reading Study in districts who use the LAS were administered Level I of both the English and Spanish versions in the Fall of each year.

Each test is administered individually. Administration time ranges from approximately 15 to 25 minutes. Trained school personnel or SEDL data collectors, who are proficient speakers of the language being tested, administered the test. The materials consist of an Examiner's Kit that includes an administration manual with stimulus pictures; an audio cassette tape of verbal stimuli for the test items; a scoring and interpretation manual; and a score sheet for each student.

The LAS incorporates a "convergent" approach to language assessment and purports to provide "an overall picture of oral linguistic proficiency based on a student's performance across four linguistic subsystems" (De Avila & Duncan, 1977, p. 1). The test consists of five subtests: Minimal Sound Pairs, Lexical, Phonemes, Sentence Comprehension, Production - Storytelling.

In the first subtest, Minimal Sound Pairs, 30 items are presented via the audio tape. For each item the student hears two words. In half of the pairs, one word differs from the other in only one phoneme (e.g., <u>very</u> / <u>berry</u>), and on the other half of the pairs, the words are the same (e.g., <u>rang</u> / <u>rang</u>). The student is asked to listen to the pair of words and to indicate verbally whether they are "the same or different." Each item is scored dichotomously as right or wrong (1 or 0, respectively) by the examiner immediately following the student's response. The student's raw score on this subtest is the number of items judged correct by the examiner.

The second subtest, Lexical, consists of 20 items. For each item the student is shown a blue and white drawing of a single object and asked to name the object in the drawing. The student's response is scored dichotomously as right or wrong (1 or 0, respectively) by the examiner at the time of testing. Credit is given for any appropriate label; probes may be used if the response given by the child is too general. The raw score for this subtest is the number of items named appropriately by the student.

In the third subtest, Phonemes, the student is presented with audiotaped stimuli for 18 pairs of items. The first item of the pair consists of a single word in which one phoneme has been isolated for scoring (e.g., <u>this</u>). The second item of the pair consists of a sentence in which the specified phoneme from the previous member of the pair exists in two of the words (My fa-th-er is fur-th-er.). As each item is presented via the audio tape, the student repeats exactly what he hears on the tape. To receive credit for the item, the student's pronunciation of the specified phoneme must be judged "correct" by the examiner each time it occurs in the stimulus. The examiner records a score of right or wrong (1 or 0, respectively) following the student's performance on each item. The raw score for this subtest is the number of items judged correct by the examiner.

For each item in the fourth subtest, Sentence Comprehension, the student is shown a series of three line drawings (pictures) arranged on a single page. While viewing the set of pictures, the student hears a sentence presented via the audio recording. The student is told to "point to the picture that shows what you heard." Credit is given if the student selects the one picture that depicts the meaning of the sentence. Each item is scored dichotomously as right or wrong (1 or 0, respectively) by the examiner at the time of administration. The raw score on this subtest is the number of items in which the student selected the "correct" picture.

In the final subtest, Production - Storytelling, the student is shown a series of four drawings and told that he will hear on the tape a story about the drawings. After hearing the story, the student is asked to retell the passage in his own words. The examiner writes down verbatim on the score sheet what the student says. Probe questions, listed on the score sheet, may be used if the student is shy or reticent. On the basis of descriptors of performance and examples of student responses included in the scoring manual, the student's performance is scored on a five-point scale (Level 1 - 5) within age group (5 years, 6-7 years, 8-9 years, 10-11 years, 12 years).

Following administration of the test, the student's raw score on each of the subtests is located on a conversion table found in the scoring manual, and the converted score is entered in the appropriate boxes at the bottom of the score sheet. A total score is then calculated based on the following equation:

$$\left[.125\left(\overset{\text{Pairs}}{\frac{\text{No. correct}}{\text{No. possible}}}+\overset{\text{Lexical}}{\frac{\text{No. correct}}{\text{No. possible}}}+\overset{\text{Phonemes}}{\frac{\text{No. correct}}{\text{No. possible}}}+\overset{\text{Comp.}}{\frac{\text{No. correct}}{\text{No. possible}}}\right)+.50\left(\overset{\text{Prod.}}{\frac{\text{Score}}{\text{Possible}}}\right)\right]\times 100=\text{Total Score}$$

Note that in this equation, the Production rating accounts for 50% of the overall score, with the remaining four subtests each contributing 12.5%. A sample of the LAS scoring sheets for the English and Spanish versions is found in Appendix B.

Reliability

For each of the language instruments employed in this study, reliability analyses (computing Cronbach's alpha) were carried out at the end of each of the five data collection years. Such analyses were performed on all data collected for a given instrument within a given year, and thus the student sample on which such analyses were based reflect the cohort structure of the study (see Volume 2: Design of the Study for a full description of this structure). In addition, given the target student replacement procedure discussed earlier (in Volume 2), a collection year sample may also reveal both withdrawn and replacement target students, dependent upon the semester in which a given instrument was administered and the time at which a given target student was replaced.

Not accounting for such target replacements, non-replaced withdrawals, or missing data for a given test administration, the target student sample by data collection year is summarized below as a review. For Year 1, the sample consisted of 40 students from Site 0 (20 at kindergarten and 20 at first grade). For Year 2, the sample contained 120 students from Sites 0-2: 50 at kindergarten, 50 at first grade, and 20 at second grade. In Year 3, the sample consisted of 380 students: 60 monolingual-Spanish first-graders (from Site 4), 40 monolingual-English kindergarteners (from Sites 3 and 5), and 280 bilingual targets (160 kindergarteners from Sites 3 and 5; and from Sites 0-2, 50 first-graders, 50 second-graders, and 20 third-graders). The Year 4 sample structure was identical to the Year 3 structure, except that each of the 380 targets was tracked into subsequent grade assignments. The Year 5 structure matched the Year 4 structure by following each student into subsequent classroom assignments, but contained only 360 students, as the 20 fourth-graders from Site 0 in Year 4 exited the study. Given this cohort structure, caution is needed in interpreting any performance difference between collection years, as the distribution of targets across grade levels and sites changes with each year. Similarly, comparisons between instruments within collection years must be made carefully due to possible differential attrition rates.

The reliability analyses for the English version of the Language Assessment Scales (LAS-E) for each collection year are presented in Table 1. These analyses were conducted only on the four multiple-item subscales (Minimal Sound Pairs, Lexical, Phonemes, and Sentence Comprehension). The first item of special note from Table 1 is that the Year 1 sample consisted of 155 students. In order to supplement the 40-student pilot sample selected in Year 1, all students in the kindergarten and first-grade classrooms from the two schools involved in the study during Year 1 were assessed with the LAS-E. Also note that the Year 3 sampl contains only 209 students, which reflects the fact that the Site 5 schools (with 110 target students) did not employ the LAS during this year (although they did use it in the subsequent years of the study). Further, none of the monolingual-Spanish students from Site 4 were assessed with this instrument during any of the data collection years.

As can be seen in Table 1, the standardized alpha coefficients computed in each collection year for three of the four subscales are all .70 or greater, indicating that the average scale score is a fairly reliable summary measure of performance on the respective task. How-

Table 1

Language Assessment Scales - English:

Reliability Analysis of the Four Multiple-Item Subscales for Each Data Collection Year

Scale	Collection Year	N of Cases	N of Items*	Item Total Mean	Item Total SD	Standardized α
Minimal Sound Pairs	1	155	30	18.7	5.6	.84
	2	127	30	20.6	6.1	.88
	3	209	30	21.3	6.1	.88
	4	306	30	23.1	4.7	.80
	5	269	30	25.3	3.3	.71
Lexical	1	155	20	10.8	5.7	.92
	2	127	20	12.7	5.9	.94
	3	209	20	15.2	4.4	.90
	4	306	20	16.6	3.3	.81
	5	269	17	15.0	2.4	.78
Phonemes	1	155	36	24.7	8.0	.91
	2	127	36	25.8	8.4	.93
	3	209	36	31.2	5.0	.88
	4	306	36	33.3	4.2	.92
	5	269	35	33.4	2.1	.[illegible]0
Sentence Comprehension	1	155	10	5.2	2.1	.56
	2	127	10	5.9	2.6	.78
	3	209	10	6.6	2.0	.59
	4	306	10	7.5	1.7	.56
	5	269	10	8.3	1.4	.38

*Items with no variance were deleted from the analysis. For each deleted item, its mean was 1.0 (i.e., all respondents answered correctly).

ever, for Sentence Comprehension, in four of the five years, the alpha coefficient falls below this level. For the four collection years where this occurs (alpha coefficients between .38 and .59), the inter-item correlations range between -.16 and .41, and since there does not appear to be a substantial floor or ceiling effect (perhaps with the exception of the Year 5 analysis), this indicates that the items contained in this subscale do not tend to reflect a single underlying dimension of performance.

In obtaining summary measures for the LAS-E, the four multiple-item subscales were averaged (without weighting for the number of subscale items) to obtain a single aggregate measure of performance. Reliability analyses over these four subscales were conducted, and the results are summarized in Table 2. The standardized alpha coefficients range from .76 to .89, suggesting that the average percent correct measure is a reliable index of performance over the four subscales. In addition to this average percent correct measure, the Production task rating and the overall LAS level were carried forward into the longitudinal analyses of growth in language skill as reflected by this instrument.

The reliability analyses of the Spanish version of the Language Assessment Scales (LAS-S) for each collection year are presented in Table 3. Again, only the four multiple-item scales were analyzed. As can be seen in the table, the pilot sample from Year 1 was again supplemented for purposes of evaluating this instrument, but rather than assessing all kindergarten and first grade students in the two Year 1 schools, only students in the four classrooms containing target students were tested. As in the LAS-E, the 110 students from Site 5 were not tested in Year 3, and further, none of the monolingual English or monolingual Spanish students were tested with the LAS-S in any collection year.

For three of the four scales, Table 3 shows that the standardized alpha coefficients in each collection year are greater than .70, indicating that the average percent correct measure for each scale is a reliable index of performance. As in the LAS-E, however, the Sentence Comprehension task possesses a reliability index below this level in Years 1 and 5 (.25 and .62, respectively). For these two years, the inter-item correlation coefficients range from -.23 to .38, again suggesting (in the absence of any floor-ceiling effects) that this scale is not measuring a single, underlying dimension of performance.

A summary index of performance on the LAS-S for the four multiple-item scales was created as in the LAS-E by averaging the four percent correct measures obtained from the scales. These four measures were entered into a reliability analysis and the results are summarized in Table 4. For all collection years except Year 1, the alpha coefficients are greater than .70, again suggesting that this averaged value reliably captures performance on the four scales. For Year 1, the alpha coefficient is .57, and inspection of the inter-item correlation coefficients shows a range from .14 to .50, with the lowest item correlations occurring with the Minimal Sound Pairs average (which

Table 2

Language Assessment Scales - English:

Reliability Analysis of the Combined Multiple-Item Subscales for Each Data Collection Year

Collection Year	N of Cases	N of Items	Item Total Mean	Item Total SD	Standardized
1	155	4	237.2	76.2	.86
2	127	4	263.3	86.7	.89
3	209	4	299.3	61.9	.83
4	306	4	327.8	47.3	.77
5	269	4	353.2	32.9	.76

197

Table 3

Language Assessment Scales - Spanish:

Reliability Analysis of the Four Multiple-Item Subscales for Each Data Collection Year

Scale	Collection Year	N of Cases	N of Items*	Item Total		Standardized
				Mean	SD	α
Minimal Sound Pairs	1	76	30	21.9	6.5	.90
	2	126	30	23.8	5.4	.89
	3	203	30	22.6	7.6	.94
	4	270	30	24.2	5.0	.86
	5	230	29	25.3	3.4	.76
Lexical	1	76	20	13.8	3.0	.74
	2	126	20	14.0	3.6	.84
	3	203	20	11.7	6.2	.94
	4	270	20	12.2	5.4	.92
	5	230	20	12.8	5.6	.93
Phonemes	1	76	31	26.8	4.6	.88
	2	126	36	31.4	5.5	.90
	3	203	36	32.2	5.3	.92
	4	270	35	32.1	4.2	.88
	5	230	32	30.0	2.8	.76
Sentence Comprehension	1	76	9	6.2	1.5	.25
	2	126	10	7.6	2.1	.75
	3	203	10	7.1	2.3	.72
	4	270	10	7.9	2.0	.72
	5	230	10	8.6	1.5	.62

*Items with no variance were deleted from the analysis. For each deleted item, its mean was 1.0 (i.e., all respondents answered correctly).

Table 4

Language Assessment Scales - Spanish:

Reliability Analysis of the Combined Multiple-Item Subscales for Each Data Collection Year

Collection Year	N of Cases	N of Items	Item Total Mean	Item Total SD	Standardized
1	76	4	302.0	42.4	.57
2	126	4	312.4	53.1	.72
3	203	4	294.5	82.0	.88
4	270	4	312.1	57.7	.73
5	230	4	332.3	50.6	.78

also shows the greatest standard deviation). As with the LAS-E, in addition to this average percent correct measure, the Production rating and the overall LAS level were carried forward into the longitudinal analyses of growth in language skill as reflected by this instrument.

Descriptive Statistics

As noted above, the summary measures created for the LAS consist of an Average Percent Correct score computed on the four multiple-item subtests, a Production score, and an overall proficiency Level rating. Recommended by the LAS test developers and in general practice in the schools, the student's language classification is assigned on the basis of the LAS Level rating; students who score at Level 3 or below are classified as Limited English (or Spanish) Proficient.

Table 5 summarizes student growth in oral language in both English and Spanish as measured by the three summary measures of the LAS. Student performance in English is reported separately from that of Spanish. For each language, performance is reported first for the total sample (overall) followed by performance of the students assigned to either Low or High proficiency groups in each language. Assignment to these groups was made on the basis of teacher ratings of the students' language abilities on the OLPRS (see below) on initial entry into the study. Based on the distribution of the total sample, students rated 3 or above (ability to participate adequately, or successfully, in school-related and peer-group conversations) in English were assigned to the High English category; those rated below 3 were assigned to the Low English category. This procedure yielded an approximately equal number of students in each category. To achieve a similar distribution for the ratings in Spanish, students rated 4 or 5 (native or near native in their ability to use Spanish) were assigned to the High Spanish category, with all others being assigned to the Low Spanish category.

For each of the summary measures, the average student growth pattern is characterized by (a) rate of growth, and (b) an estimate of where the average student was on entry into kindergarten (see Volume 3 for a discussion of the growth measures). In Table 5, the LAS summary measures appear along the left margin; all English measures are listed first, followed by the Spanish measures. Two growth measures (Intercept and Slope) for each summary measure are provided. Three statistics (mean, standard deviation, number of students) is shown for each of the growth measures.

LAS-English. As can be noted in Table 5, the estimated average student entry level rating in English (LASLE; S-Intrcp) at kindergarten is below Level 3 for the overall sample as well as for each category of students. On the basis of this measure, this group of students, as a whole, would be classified as Limited English Proficient. As indicated by the Slope, the rate of growth is similar across groupings, with students gaining about three-quarters of a level per year. The slowest growth (0.6) is shown by the students in the High Spanish category.

Table 5

Language Assessment Scales
Descriptive Statistics on Growth Measures from the Entire Sample and for Each Language Entry Group

Instrument	Measure	Statistic	Overall	Low Eng.	High Eng.	Low Span.	High Span.
LASLE	S-Intrcp	M	2.1	1.4	2.8	2.1	2.2
LASLE	S-Intrcp	S	1.7	1.5	1.6	1.5	1.8
LASLE	S-Intrcp	N	254	116	138	113	141
LASLE	Slope	M	0.7	0.8	0.7	0.8	0.6
LASLE	Slope	S	0.8	0.8	0.7	0.7	0.8
LASLE	Slope	N	254	116	138	113	141
LASAE	S-Intrcp	M	66.9	57.1	75.2	66.9	66.9
LASAE	S-Intrcp	S	17.5	17.5	12.8	17.6	17.6
LASAE	S-Intrcp	N	254	116	138	113	141
LASAE	Slope	M	8.5	10.6	6.8	9.4	7.8
LASAE	Slope	S	6.0	6.4	5.1	6.1	5.9
LASAE	Slope	N	254	116	138	113	141
LASPE	S-Intrcp	M	2.3	1.6	2.9	2.3	2.3
LASPE	S-Intrcp	S	1.6	1.6	1.4	1.4	1.8
LASPE	S-Intrcp	N	254	116	138	113	141
LASPE	Slope	M	0.5	0.5	0.5	0.5	0.5
LASPE	Slope	S	0.7	0.8	0.7	0.7	0.8
LASPE	Slope	N	254	116	138	113	141
LASLS	S-Intrcp	M	2.4	2.6	2.3	1.8	2.9
LASLS	S-Intrcp	S	1.7	1.7	1.6	1.5	1.6
LASLS	S-Intrcp	N	254	116	138	113	141
LASLS	Slope	M	0.2	0.2	0.2	.0	0.3
LASLS	Slope	S	0.8	0.8	0.8	0.8	0.8
LASLS	Slope	N	254	116	138	113	141
LASAS	S-Intrcp	M	68.5	68.1	68.9	60.5	74.9
LASAS	S-Intrcp	S	16.2	16.5	16.0	15.3	13.9
LASAS	S-Intrcp	N	254	116	138	113	141
LASAS	Slope	M	6.6	7.2	6.1	7.7	5.7
LASAS	Slope	S	5.9	5.5	6.3	6.0	5.7
LASAS	Slope	N	254	116	138	113	141
LASPS	S-Intrcp	M	2.4	2.7	2.2	1.8	3.0
LASPS	S-Intrcp	S	1.8	1.9	1.8	1.8	1.7
LASPS	S-Intrcp	N	254	116	138	113	141
LASPS	Slope	M	0.1	.0	0.2	.0	0.1
LASPS	Slope	S	0.9	0.9	0.9	0.9	0.9
LASPS	Slope	N	254	116	138	113	141

To get a sense of the students' ability to deal with the kinds of tasks measured by the LAS, let us next consider the two summary measures from which the Level rating is derived. On the multiple-item subtests (LASAE) the students showed considerable skill; they were able to respond correctly to approximately 60% to 75% of the items and demonstrated a growth rate of approximately 10% each year. On the Production task, however, the students were less successful. They entered at about a Level 2 (the students in the High English category were close to a Level 3) and gained only about a half a level per year.

Data on the performance on the LAS-English by the overall sample is presented graphically in Figure 1.

To summarize briefly, the average student in this sample entered kindergarten with the ability to handle discrete items fairly well in English and took about three and one-half years (mid-third grade) to reach mastery on those items. The Production task, on the other hand, presented a more difficult challenge. The students in general were less able to comprehend and restructure narrative text. With schooling and literacy instruction, they grew at about a half level per year but were projected to exit 4th grade at slightly below Level 5. The LAS Level rating in the early years is strongly affected by the Production score. As the students gain some skill in this task and function at a high level on the discrete-item subtests, the LAS Level rating places the students at Level 5 at about mid-third grade.

<u>LAS-Spanish</u>. As noted in Table 5, the LAS Level entry scores (LASLS; S-Intrcp), while slightly higher in Spanish than in English, are nonetheless below Level 3 for all categories of students. The growth rates are somewhat lower in Spanish than in English, with the students in the High Spanish and Low English categories having the highest entry scores as well as the highest growth rates.

Student performance on the multiple-item subtests is slightly higher in Spanish than in English for the overall sample and is above 60% for all groups of students. Their growth rate on these tasks is noticeably lower in Spanish, however.

On the Production task, the entry scores are about the same in both languages for the overall sample but are lower in Spanish for all categories of students, except for those in the Low English category. The growth rates for all groups of students are extremely low.

The data in Figure 2 summarizes student LAS performance in Spanish. It is not at all the picture one might expect. The students, on entry into school, appear to have slightly greater strength in Spanish than in English but show little growth in their ability to use Spanish to accomplish the tasks measured by the LAS. In contrast, their ability to perform these tasks in English showed considerable growth during the years of the study. It should be noted, however, that as a group their scores on the Production task in both Spanish and English suggest that comprehending and restructuring connected text presents a challenge for these students, particularly in the early grades.

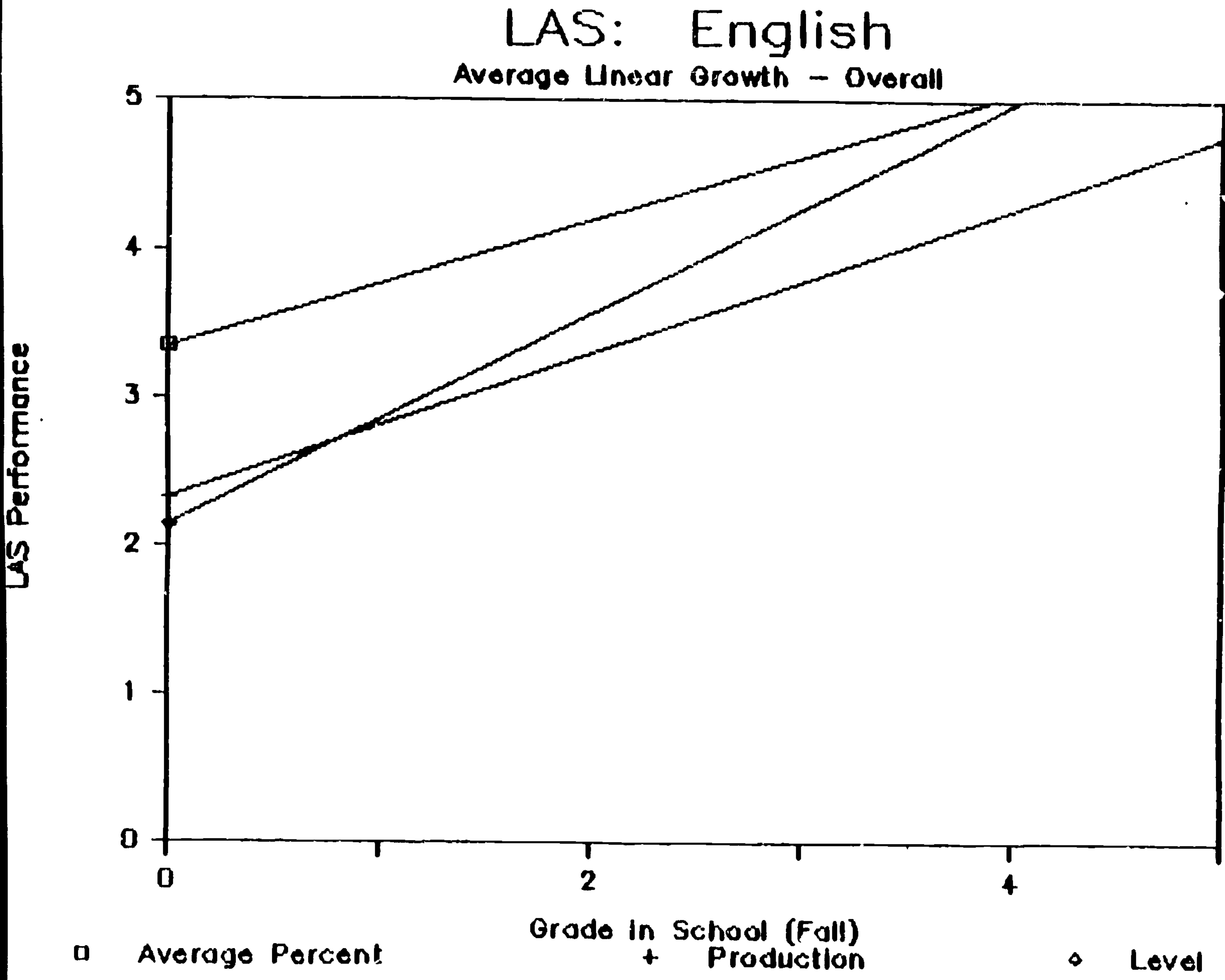


Figure 1. Average linear growth for three LAS (English) summary measures over entire sample.

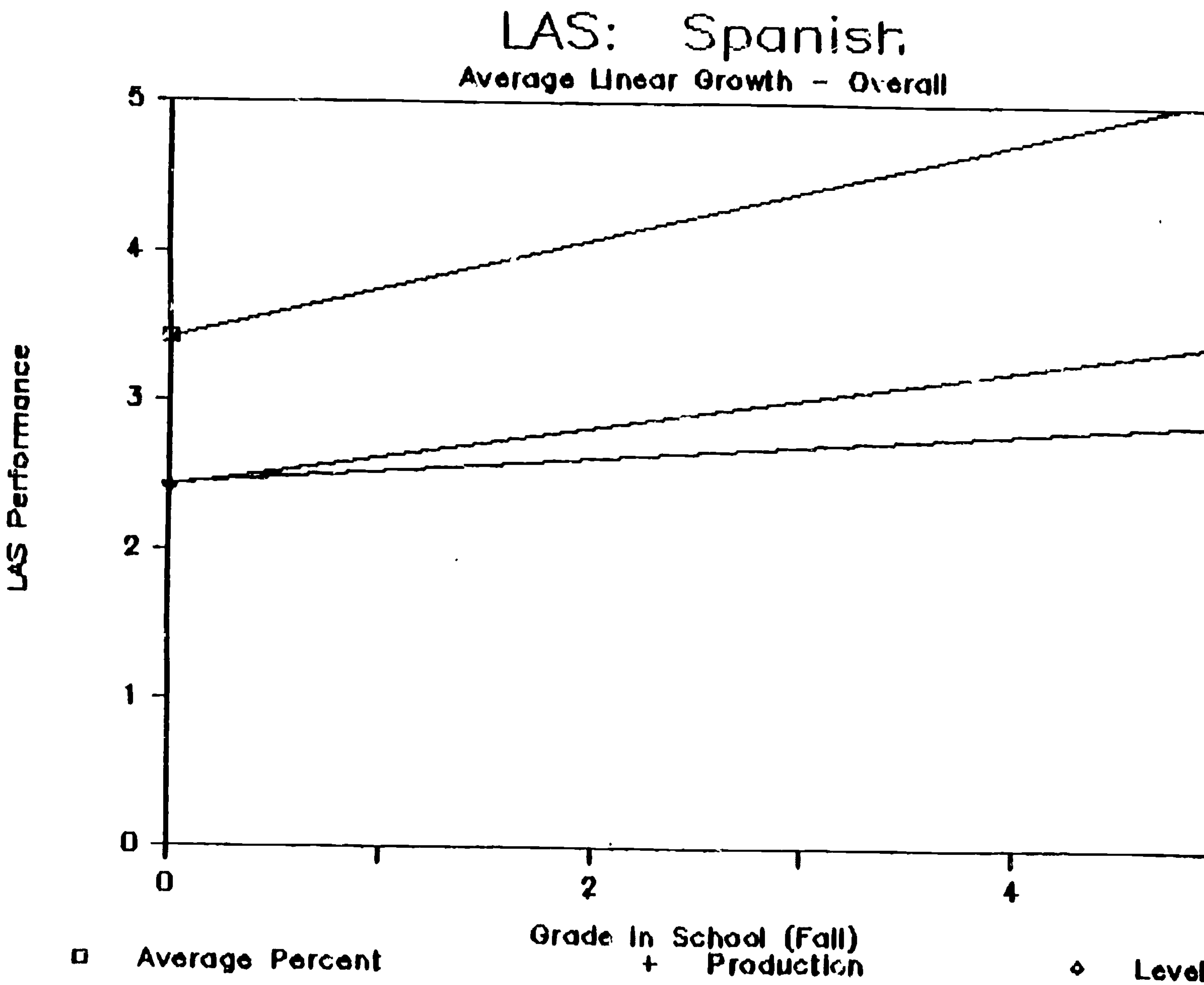


Figure 2. Average linear growth for three LAS (Spanish) summary measures over entire sample.

Student performance on the LAS by site. As can be seen in Figures 3 and 4, the LAS Level scores suggest site differences both at entry and as the students progress through the grades. The students entered with differing levels of abilities in English, but showed similar growth rates at four of the sites. The students at Site 2, however, entered with greater skill in English than did the others but their growth rates were lower.

The sites differed considerably in student performance in Spanish. Both entry scores and growth rates show wide variation. Only two of the sites (Site 0 and Site 1) show substantial growth in Spanish. Both of these sites are rural, located along th Texas-Mexico border, and maintain close ties with Mexico. Site 2 is also rural and located near the border, but the children in this district have greater exposure to English both in the local community and in the city 35 miles away where a number of the parents commute to work. The children in Site 3 tend to be English dominant on entry into school and maintain Spanish at a low level. Site 5 is located in a large urban area. The students at this site appear to maintain their Spanish but to show little growth in it over time.

Oral Language Proficiency Rating Scale

Teacher observation and rating of student language performance was used in the study, along with other measures, to develop an index of the student's oral language ability. Teachers, generally, have at best minimal training in language assessment and are not consciously aware of how sociocultural variables influence the manner in which morphological, phonological, and lexical items are integrated into cohesive discourse (Rivera & Simich, 1982). They do, however, have a working knowledge of traditional linguistic terms such as pronunciation, grammar, vocabulary, and comprehension. Recognizing these constraints and, at the same time, recognizing that teachers are better qualified to make valid predictions about their own students' language abilities than are outsiders (who generally would not be aware of the specific rules of interaction implicitly or explicitly agreed upon by participants in classrooms settings), the research staff at SEDL developed the Oral Language Proficiency Rating Scale - OLPRS (Mace-Matluck, Tunmer, & Domínguez, 1979).

Tasks, Materials, Scoring

To provide a familiar framework within which to approach observation, and to get teachers to consciously focus their attention on student's language performance, the teachers were asked to observe the language performance of particular students in their classes. They were then asked to rate, on the basis of a set of descriptors, the language performance of their students on a five-point scale (1 to 5, from lowest to highest rating) for each of four language components (pronunciation, grammar, vocabulary, comprehension) and for a fifth category identified as "Overall Communicative Skill." This fifth category resulted in a holistic rating (also based on a set of descriptors) of each student's ability to participate in school-related or peer-group

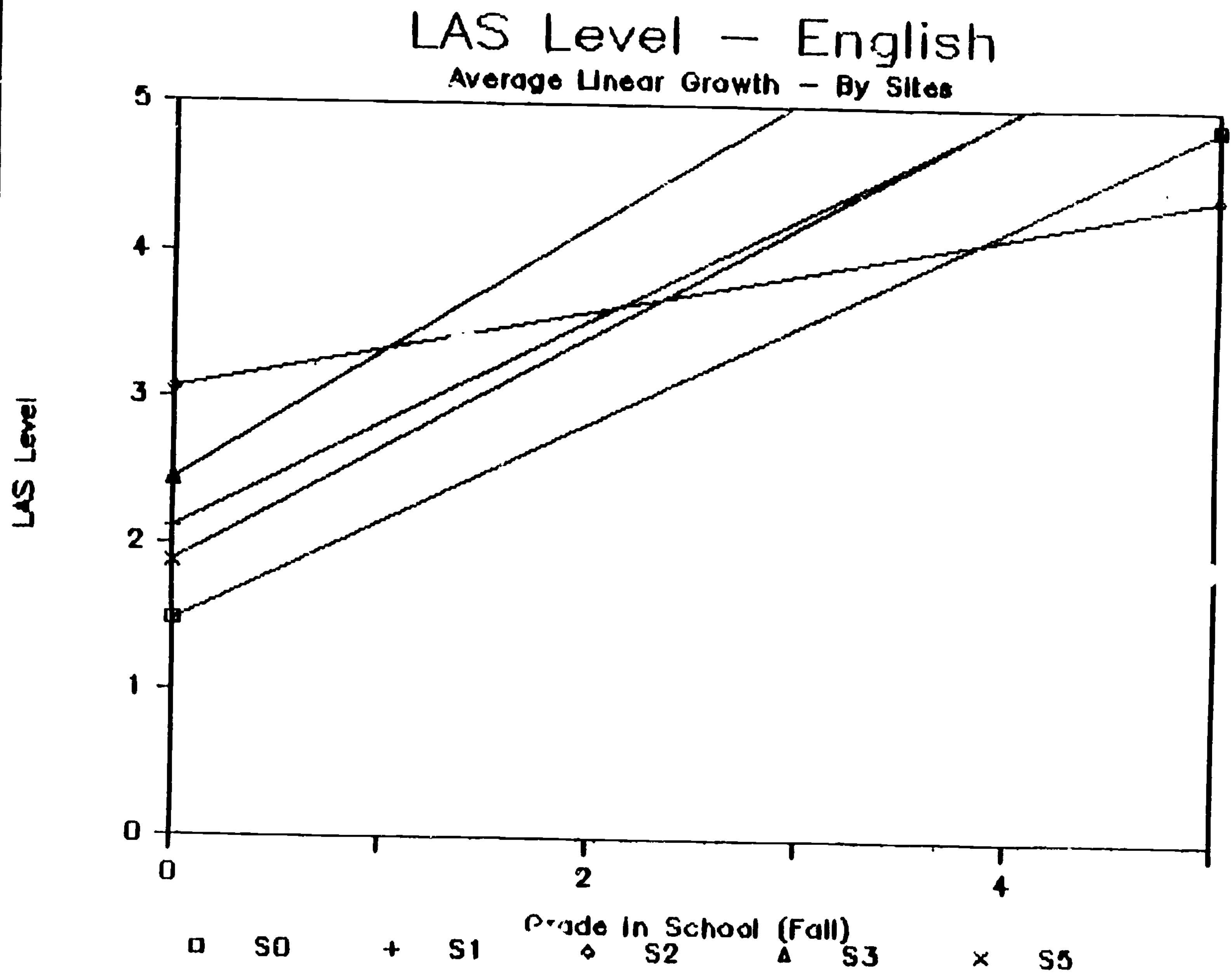


Figure 3. Average linear growth for LAS le... (English) measure for entire sample by site.

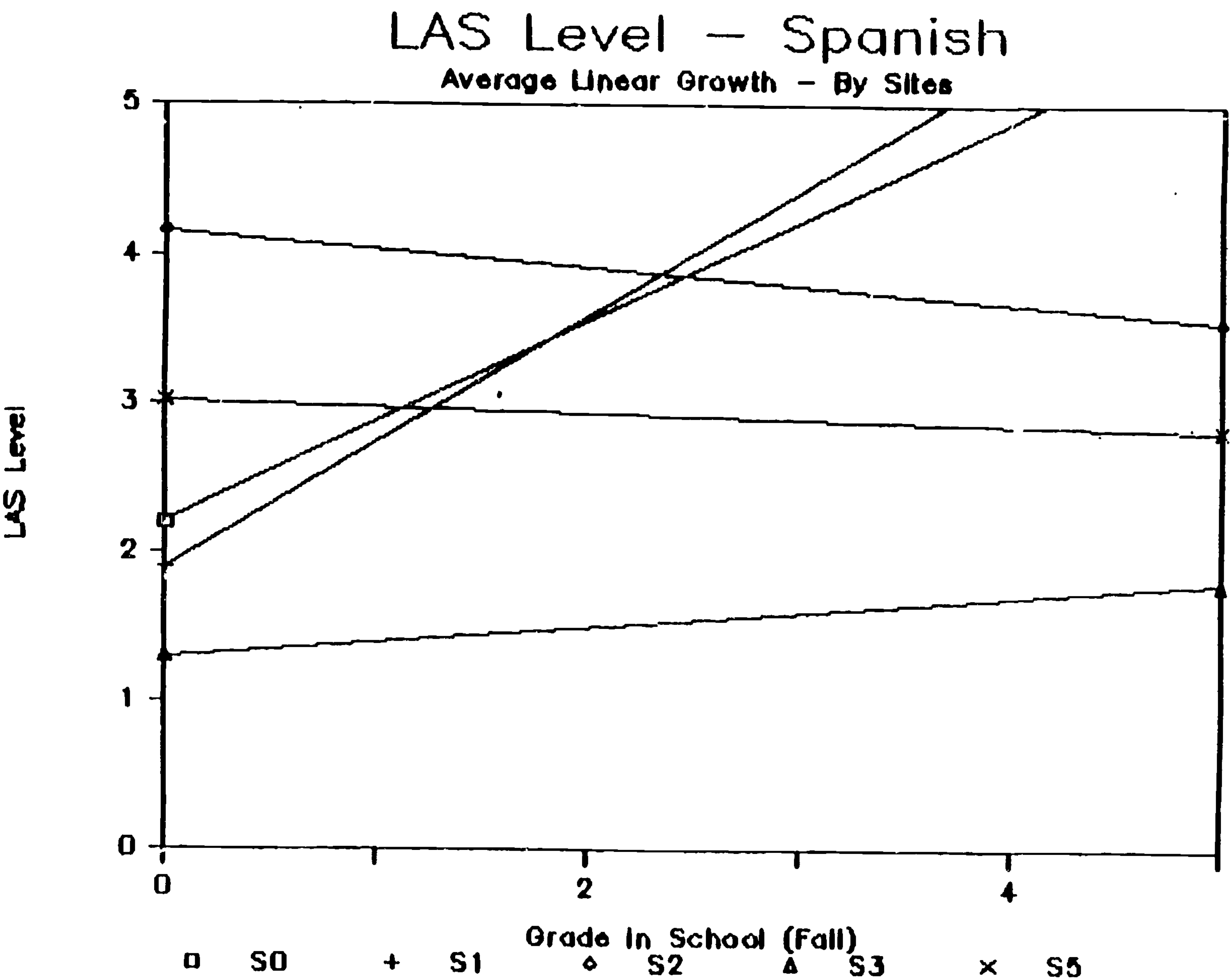


Figure 4. Average linear growth for LAS level (Spanish) measure for entire sample by site.

interactions in a given language within the school setting. It is the rating in the fifth category, overall communicative skill, that was of particular interest to the research team. A copy of the instrument is provided in Appendix C.

OLPRS ratings were obtained for each student, for both English and Spanish, twice during each school year. The first was obtained in the month of December, after teachers had become familiar with the language patterns and usage of their students; the second rating was made in April, concurrent with the administration of the reading achievement measures. Summary measures of performance were generated by averaging the 5 subscale scores to obtain a single index of the teacher's ratings of each student's proficiency in each language for each semester.

Reliability

Cronbach's alpha was computed for each of the sets of rating data (English and Spanish; Fall and Spring semester) for each data collection year. Again, the cohort structure of the sampling plan is reflected in these analyses as explained above. The results of these analyses for the English data sets are summarized in Table 6.

As can be seen, no teacher ratings were obtained for the Year 1 target students in the Fall, but all were rated during the Spring semester. In Year 3, the 20 student difference between Fall and Spring ratings reflects target student replacement (kindergarteners withdrawing in Sites 3 and 5 in late Fall just prior to the December ratings, with their replacements not being selected until early Spring). The discrepancy in number between Fall and Spring ratings in Year 4 was due to a data collection oversight in some classrooms at Site 3; a similar error was made in Sites 0 and 2 in the Fall of Year 5.

Table 6 shows that in each of the 5 collection years and for each semester, the standardized coefficient alpha was .96 or greater, indicating that the average score across the 5 scales is a highly reliable indicator of the teacher's ratings. Thus, although the overall communicative skill rating was of primary interest in this instrument, ratings from the other 4 scales were strongly related to it.

Table 7 presents the results of a similar reliability analysis conducted on the Spanish ratings (by collection year and semester). In addition to the explanations provided above for discrepancies in sample size, there is a substantial decrease in the number of Spanish ratings obtained relative to the number of English ratings. This occurred as the students advanced in grade level and left bilingual instruction for exclusive English instruction, and thus, mainly monolingual-English teachers, who were not capable of rating their students' Spanish oral language skills, and were asked to make the ratings in English only.

As with the English ratings, Table 7 shows that each of the standardized alpha coefficients is high (.95 or greater), again indicating that the scale average across the 5 rating scales is a highly reliable index of the teachers' ratings of Spanish oral proficiency.

Table 6

Oral Language Proficiency Rating Scale - English:

Reliability Analysis of the Total Scale Score for Each Semester for Each Data Collection Year

Semester	Collection Year	N of Cases	N of Items	Item Total Mean	Item Total SD	Standardized
Fall	1	0	-	-	-	-
	2	118	5	15.4	7.6	.99
	3	289	5	16.9	6.5	.98
	4	304	5	17.8	5.9	.98
	5	217	5	18.7	5.4	.98
Spring	1	41	5	15.5	5.2	.96
	2	118	5	16.6	6.4	.98
	3	309	5	18.3	5.6	.97
	4	270	5	18.8	5.8	.98
	5	262	5	20.0	5.0	.97

209

Table 7

Oral Language Proficiency Rating Scale - Spanish:

Reliability Analysis of the Total Scale Score for Each Semester for Each Data Collection Year

Semester	Collection Year	N of Cases	N of Items	Item Total Mean	Item Total SD	Standardized
Fall	1	0	-	-	-	-
	2	109	5	22.2	4.5	.95
	3	229	5	17.0	6.4	.96
	4	229	5	17.8	5.4	.97
	5	149	5	17.1	5.9	.98
Spring	1	41	5	21.4	4.6	.96
	2	104	5	22.3	3.6	.96
	3	278	5	18.0	6.5	.97
	4	230	5	18.6	5.9	.97
	5	155	5	19.5	5.2	.97

210

Descriptive Statistics

The summary measure for the OLPRS is a single index of the teacher's rating of each student's proficiency in each of the language for each semester. The measure was generated by averaging the five subscale scores in each rating.

In Table 8, are the basic descriptive findings for the OLPRS measure. Presented first are the data for the Fall (OLPRSFE) and Spring (OLPRSSE) ratings of the students performance in English. Following these are the the Fall and Spring (OLPRSFS; OLPRSSS) ratings for Spanish.

OLPRS-English. The teachers' ratings indicate that on the average the students in the sample were able to participate adequately or successfully in school-related and peer-group conversations (Level 3) relative to their grade level on entry into the study. Two groups, Low English and Low Spanish, received ratings that indicate that they were unable to participate fully in school-related or peer-group activities conducted in English. The Spring ratings, however, are somewhat higher for these students such that only those students in the LOW English category received a rating below Level 3. The growth rates are low for all categories of students for both the Fall and Spring ratings. This, we believe, does not suggest that students are not gaining in their English language skills, but rather that as the demands of the classroom increase the relationship between the students language skills and their ability to handle instruction in English at the higher grade level is taken into consideration when the teachers made their ratings. This, in fact, was confirmed informally in conversations with the teachers.

Figures 5 and 6 show graphically the ratings of students discussed above. In the Fall ratings, students who were rated lowest at entry made the greatest gains. Students who were rated High English or High Spanish appeared to either loose ground or progress only minimally, in relation to the demands of the classroom from year to year. The Spring ratings show all categories of students making slight gains in their English skills. Note that the Spring entry ratings are somewhat higher than the Fall ratings for the Low English and Low Spanish category students, reflecting considerable growth during a given school year.

OLPRS-Spanish. As can be seen in Table 8, the general picture that emerges from both the Fall and Spring ratings is that the teachers see this group of students as entering school with native or near-native ability to use Spanish but to show little growth in Spanish, in relation to the demands of the classroom at subsequent grade levels. The one exception is the students in the Low Spanish category who show considerable gains in the Fall ratings over time. A comparison of the entry Fall ratings with the entry Spring ratings suggests that this group of students noticeably improve their Spanish skills in their initial year of schooling. Data from Table 5 on student performance in Spanish is displayed in Figures 7 and 8.

Table 8

Oral Language Proficiency Rating Scale:
Descriptive Statistics on Growth Measures from the Entire Sample and for Each Language Entry Group

Instrument	Measure	Statistic	Overall	Low Eng.	High Eng.	Low Span.	High Span.
OPRSFE	S-Intrcp	M	3.1	1.7	4.2	2.8	3.2
OPRSFE	S-Intrcp	S	1.5	0.9	0.9	1.4	1.6
OPRSFE	S-Intrcp	N	251	114	137	112	139
OPRSFE	Slope	M	0.2	0.6	-0.1	0.4	0.1
OPRSFE	Slope	S	0.6	0.6	0.5	0.6	0.7
OPRSFE	Slope	N	251	114	137	112	139
OPRSSE	S-Intrcp	M	3.3	2.5	4.0	3.3	3.3
OPRSSE	S-Intrcp	S	1.4	1.2	1.2	1.3	1.4
OPRSSE	S-Intrcp	N	253	116	137	112	141
OPRSSE	Slope	M	0.2	0.2	0.1	0.2	0.1
OPRSSE	Slope	S	0.5	0.5	0.5	0.5	0.4
OPRSSE	Slope	N	253	116	137	112	141
OPRSFS	S-Intrcp	M	3.8	3.5	4.1	2.3	5.0
OPRSFS	S-Intrcp	S	1.7	1.8	1.6	1.2	1.0
OPRSFS	S-Intrcp	N	222	108	114	96	126
OPRSFS	Slope	M	-0.1	0.1	-0.3	0.5	-0.5
OPRSFS	Slope	S	0.9	0.8	0.9	0.8	0.6
OPRSFS	Slope	N	222	108	114	96	126
OPRSSS	S-Intrcp	M	4.0	4.1	3.9	2.9	4.8
OPRSSS	S-Intrcp	S	1.6	1.6	1.7	1.7	1.0
OPRSSS	S-Intrcp	N	203	100	103	84	119
OPRSSS	Slope	M	-0.1	-0.1	.0	0.1	-0.2
OPRSSS	Slope	S	0.6	0.5	0.7	0.8	0.4
OPRSSS	Slope	N	203	100	103	84	119

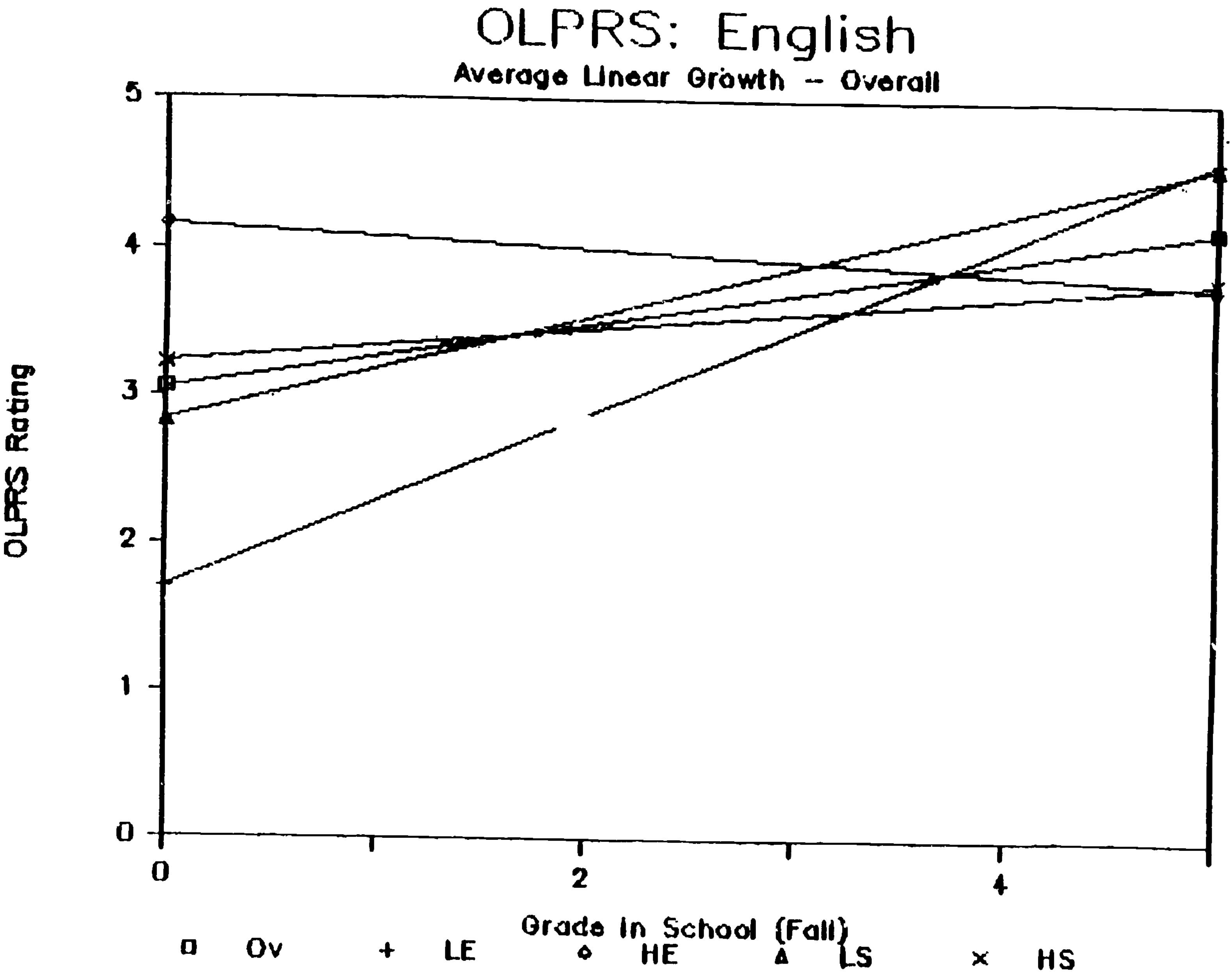


Figure 5. Average linear growth for Fall OLPRS (English) over entire sample and for each language entry category.

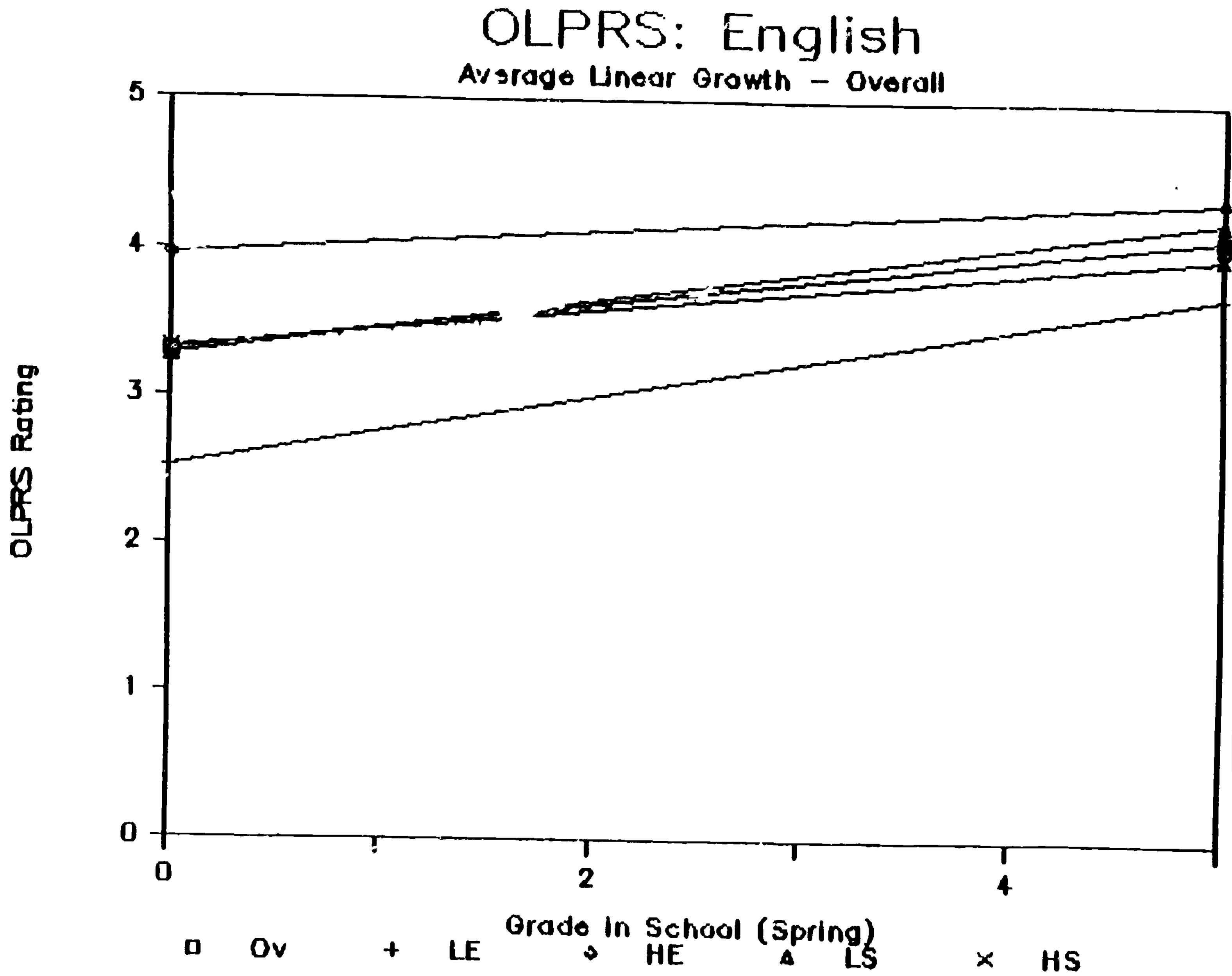


Figure 6. Average linear growth for Spring OLPRS (English) over entire sample and for each language entry category.

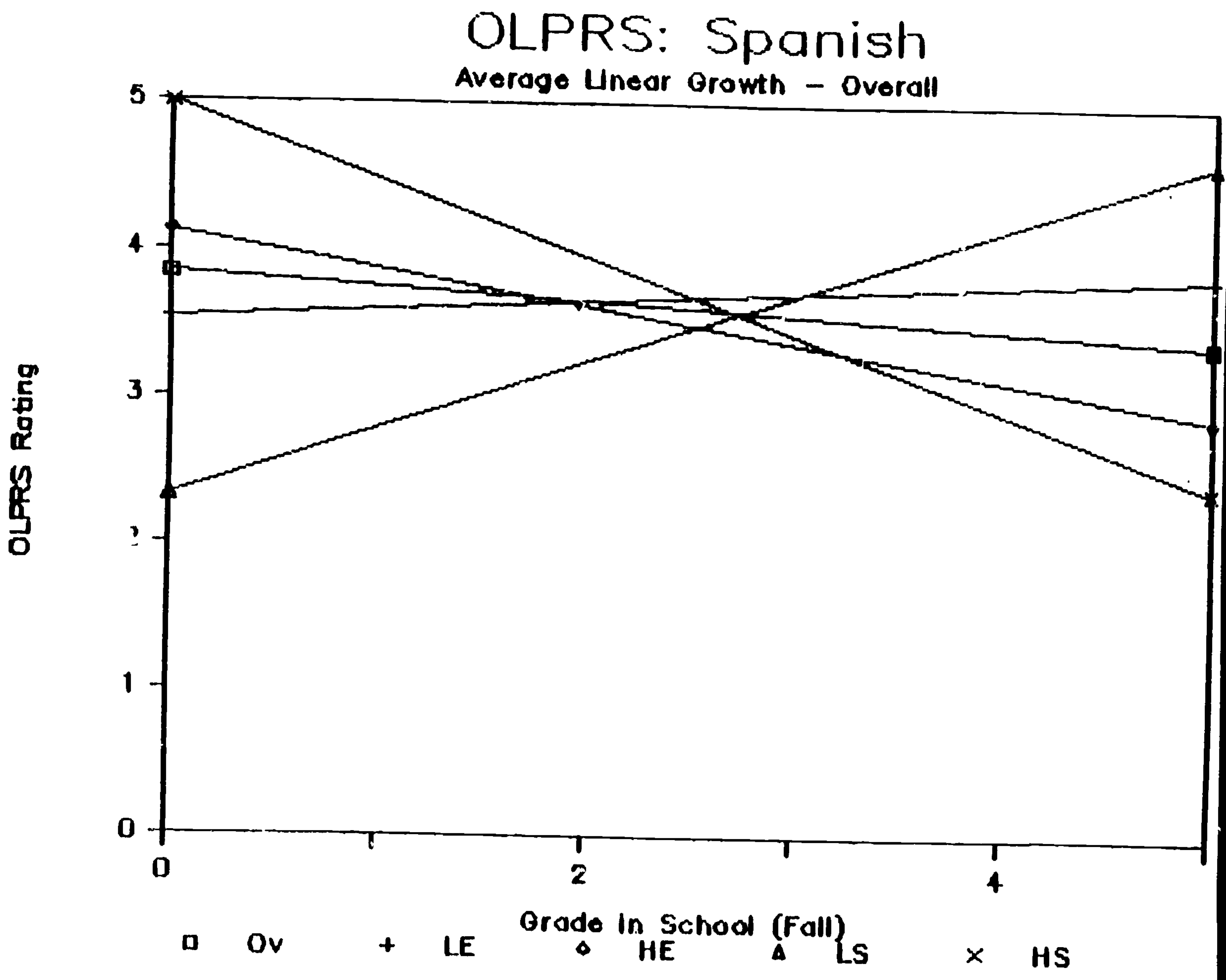


Figure 7. Average linear growth for Fall OLPRS (Spanish) over entire sample and for each language entry category.

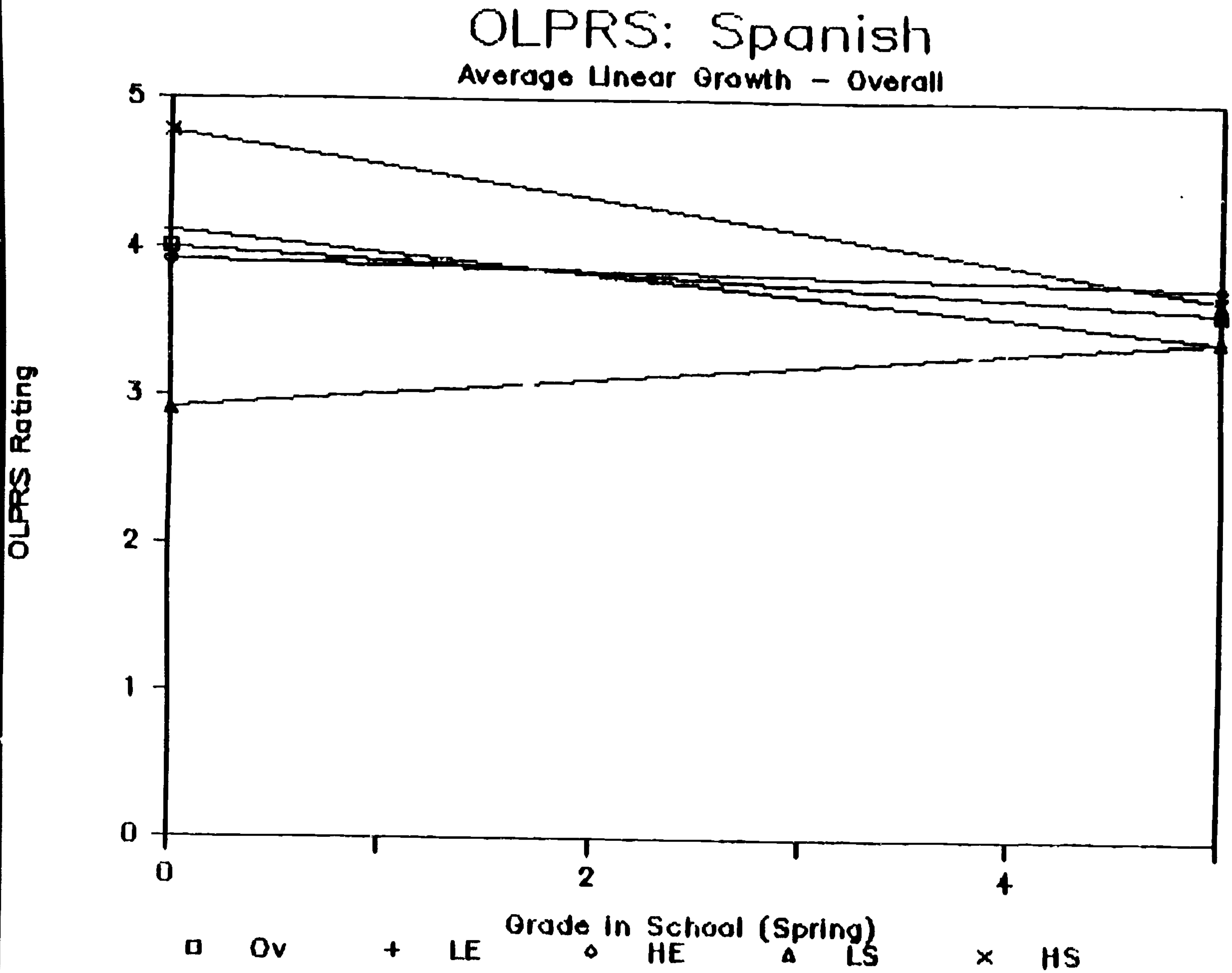


Figure 8. Average linear growth for Spring OLPRS (Spanish) over entire sample and for each language entry category.

Student performance on the OLPRS by site. The language environment of the students from which the sample was drawn differed considerably. Students in the border sites (Site 0, Site 1, and Site 2) were in schools where the student population was essentially Hispanic, and from homes where Spanish was spoken; Spanish was also widely spoken in the community. In Site 3, in the Central Texas area, the students heard Spanish spoken in their homes, but English was the primary language of the wider community. The students in Site 5 live in an urban area, but their homes, generally, are located in neighborhoods where Spanish is widely spoken. Thus, the children in Sites 3 and 5 are much more likely to have wider exposure to English on entry into school and less exposure to Spanish than are the students in the three border sites. Site differences are apparent in the OLPRS ratings of these students both on entry and in subsequent years.

As can be noted in Figure 9, Fall entry ratings for the students in English at all sites is quite similar (at about mid-scale). Growth in English is lower for students in Site 0 and Site 1 than for those at the other three sites. Site 2, while located near the border, differs from the other two border sites in its relationship to English-speaking communities. This situation is commented upon further in a subsequent section of this Volume. The Spring ratings for the sites show similar patterns.

The Fall entry ratings in Spanish (Figure 10) reflect the children's linguistic environment. All groups, except at Site 3, were rated as native or near-native in Spanish on entry into school. The students at Site 5 maintained their entry level rating over time, in relation to the demands of the classroom, and those of Site 3 showed a small increase over their entry level rating. It would appear that the students at these two sites were growing in their ability to use the formal aspects of Spanish. In the border area, students in Sites 1 and 2 show some decline in their ability to perform in Spanish in relation to classroom demands, but still are estimated to be above mid-scale at the end of fourth grade. The students at Site 0 present a very different picture. They enter with native-like ability in Spanish but show a sharp decline in their ratings over time. The Spring ratings (Figure 11) show a very interesting pattern. During the students' initial year in the study (kindergarten in most cases), their Spanish skills increase from Fall to Spring, except for Sites 0 and 3, where some decline is shown. Their exit scores, based on the Spring ratings at the end of the initial year, suggest that there is a general decline in their ability to function in Spanish in the classroom over time (except for Site 3 where growth is shown) and that the students at Site 0 are also projected at above mid-scale at the end f fourth grade, with less of a decline than shown in the Fall ratings. Interpretation of these data is difficult to, to say the least. However, both sets of data (Fall and Spring ratings in Spanish) suggest that, except for Site 3, the students' Spanish skills, in the eyes of their teachers, are less adequate for school purposes as they progress in school than they are on entry.

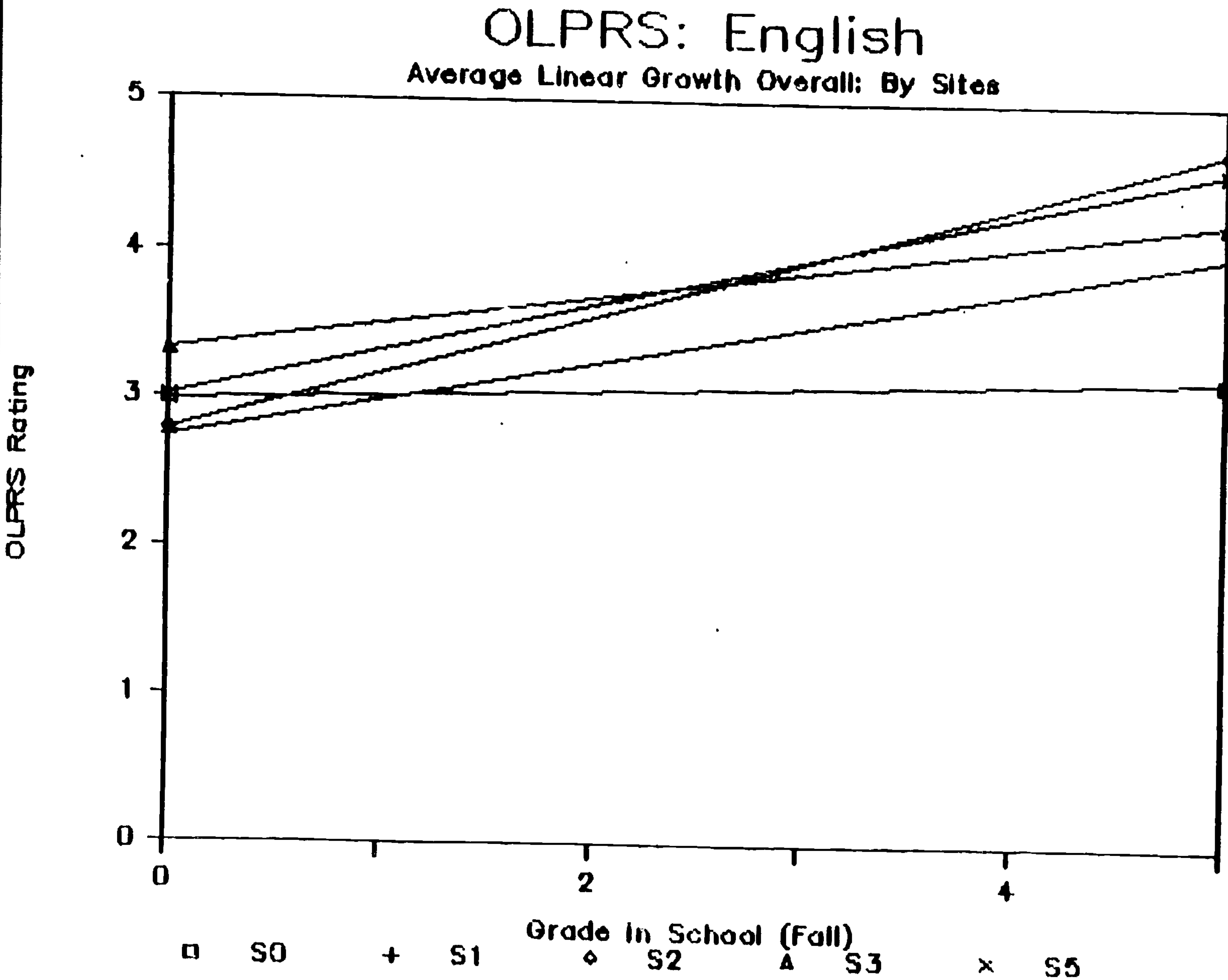


Figure 9. Average linear growth for Fall OLPRS (English) by sites.

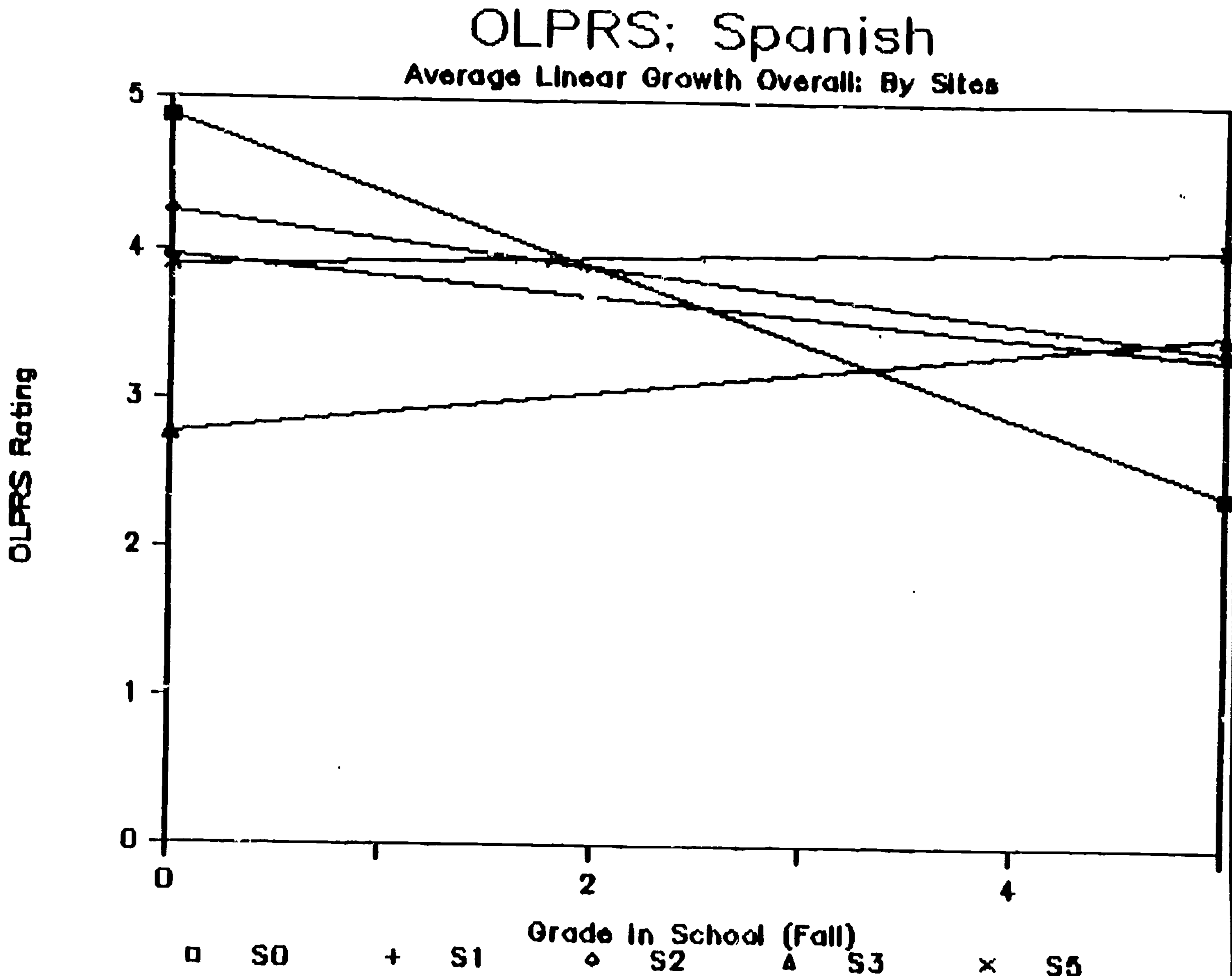


Figure 10. Average linear growth for Fall OLPRS (Spanish) by sites.

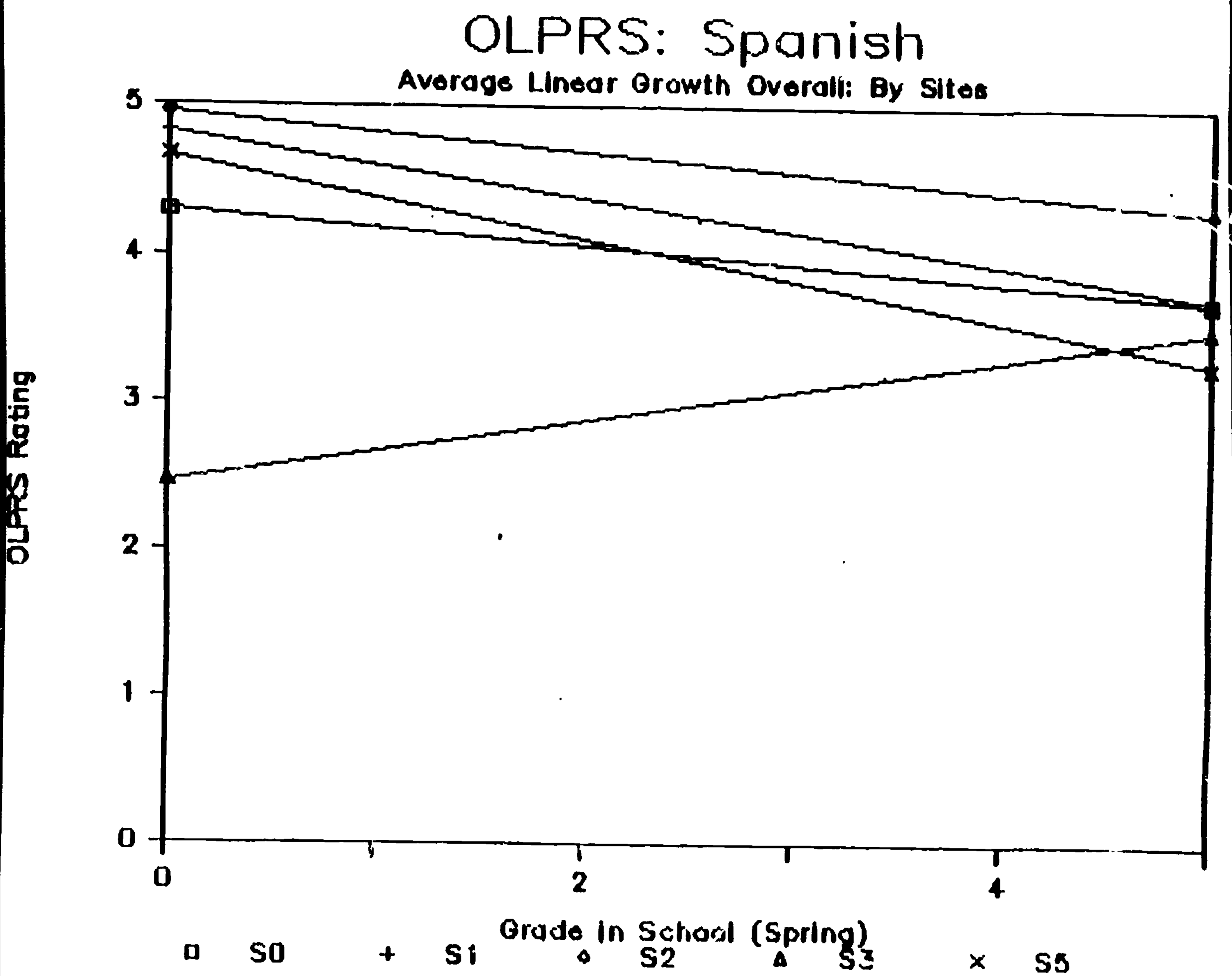


Figure 11. Average linear growth for Spring OLPRS (Spanish) by sites.

Taped Interactions - Language Samples

A number of scholars studying the relationship between language and thought have discussed the use and interpretation of language in different contexts; others have noted the effects on language behavior of differing participants (e.g., age, sex, status relationships) and differing topics. In an effort to gain as wide a representation of the student's language abilities, for the purpose of monitoring the student's language growth as well as verifying information obtained on other measures, audiotaped speech samples were taken once a month from selected target students. The samples were taken on a rotating schedule in three communication settings: in the classroom, in the home, and either on the playground or in other noninstructional settings within the school. Procedures for obtaining and evaluating these samples were developed by the SEDL research staff (Mace-Matluck, Tunmer, & Domínguez, 1978).

The taped interactions have provided the study with a rich data bank from which valuable insights have been gained into the language development of the students, their language preference in each of the communication settings, and the patterns of language use that are found within the student's environment.

Tasks, Materials, Scoring

The taped samples for each child are twenty-to-thirty minutes in length. For taping in the classroom, standard, high-quality cassette tape recorders and lapel microphones were used. Generally, the data collector, after ensuring that the teacher was familiar with the recording equipment, left the room and returned at an appointed time to collect the e pment and tape. The teachers were instructed to tape instructional segments or typical classroom interactions which involved the target children. In many cases, those selected by the teachers for taping were small-group lessons from the various content areas. Others consisted of the teacher interacting with a single student either in an instructional role or as a conversation partner. The latte were more prevalent in the data from the kindergarten and first grad lasses.

The taped samples on the playground and in the home were obtained by placing an activated microcassette tape recorder (Sony M-102B) in the pocket of a specially-designed belt-and-sash worn by the child. This is similar to that worn by children on school-crossing patrol. A very small lapel microphone extended from the tape recorder up under the sash and through a buttonhole at shoulder height, ensuring that the microphone was ideally placed to pick up the student's speech, as well as that of others around him. After recording the identifying information and potential interlocuters and ensuring that the equipment was recording, the data collector withdrew from the scene, although in the case of the playground setting, the data collector remained nearby to be of assistance if needed. The data collector also engaged the target child in a brief conversation and recorded this interaction on the beginning footage of the tape to assist the researchers in identifying the voice of the target student from among his peers. The placement of the microphone also assisted in this matter since the volume and quality of the target student's voice differed from that of the students who were not wearing the equipment.

During the first two years of the study all target children (N = 40 in Year 1, and N = 120 in Year 2) were taped once a month. However, such extensive taping heavily taxed the fiscal and human resources available to the study; and, after examining the language data collected during those two years, it was decided that the validity of the oral language proficiency test scores and teacher ratings could be adequately evaluated via tape data obtained from a reduced sample of 60% of the original sample. Thus, only 6 students from the originally selected 10 targets per classroom were taped in subsequent years. The taping schedule was further revised the following year to limit the taping in the home to one session per year, as opposed to the two that were scheduled for the first two years.

For the first three years of the study, each of the tapes were transcribed by a bilingual speaker of English and Spanish. In subsequent years, because of limited resources, only selected tapes were transcribed. SEDL staff members and/or university graduate students, who have expertise in oral language assessment and linguistics, examined the taped samples (and the transcripts where available) for extent and quality of language and for language preference in each of the communication settings. As the tapes were evaluated, certain information was recorded about the total interaction (e.g., general language of the student and of the interlocuters, dialect variations, instances of codeswitching and language alternation, errors in phonological and grammatical structures). After extensive examination of the language sample, the student was assigned a rating in each of the categories of the OLPRS (a 1 to 5 point scale from lowest to highest, as described above), using the same criteria used by the teachers in making their ratings on the OLPRS. A sample of the rating form is provided in Appendix D.

As sometimes happens when relying on natural, or spontaneous, speech samples, the data are inadequate for the purposes intended (e.g., the target student says very little during the interactions, joins his friends to watch a program on television, or the tape recorder malfunctions). To be rated and included in the analyses, the sample had to be adequate, both in quantity of speech and quality of the recording, to allow a judgment to be made about the student's performance in each of the categories which comprised the rating. The speech samples, generally, were of good quality and contained sufficient participation by the target student to be rated in one or both languages.

In deriving a summary measure for the tape ratings, for each tape available in each setting where there was a sufficient sample to rate, and for each language rated on such tapes, the 5 OLPRS rating scale values were averaged to obtain an overall scale measure. Since multiple ratings were available within settings for most students, a comparison between the largest average rating and the smallest average rating was made within each setting and language. The results of this comparison are presented in Table 9. Over the 2,328 rateable samples collected over the 5 year data collection phase of the study, less than 1% show average rating differences which exceed 1 point on the five-point rating scale, and less than 5% show an average difference greater than 0.5. It is important to remember that this comparison is made within settings and languages, and speaks only to the comparability of multiple samples within a language-setting combination--it does not suggest that the tape ratings show little difference between settings or over time.

Given that the differences between average ratings within a given setting and language were found to be minimal, any of the available averages could have been selected to represent a given student's proficiency rating within a setting without introducing much error. However, in order to be systematic, and to give the student the "benefit of the doubt," the maximum average was selected to represent performance whenever multiple ratings were available for a given setting within a given year.

In order to assess whether there were significant differences between ratings when multiple-setting ratings were available, comparisons between pairs of such ratings were made. The analysis was conducted only on pairs of available English ratings, and only for the bilingual sample, since the average tape rating summary measures were generally at the top of the scale for (1) English usage by the monolingual-English sample, and (2) Spanish usage by both the monolingual-Spanish and bilingual samples alike. Within each instructional year, the maximum average within each setting for each student was selected, and comparisons of these maximum averages were made between settings. Descriptive statistics on the maximum average for each member of a setting pair (Classroom-Playground, Classroom-Home, and Playground-Home) in each instructional year are presented in Table 10. From the table, there appears to be little systematic difference between maximum average rating means (or standard deviations) between setting pairs. Given this, for each language, the maximum average rating across settings was selected to represent each student's oral proficiency as indexed by the tape ratings.

A second important piece of information from this data set concerns language preference within the three settings. Descriptive data on such preferences are discussed in a subsequent section of this document.

Reliability

For each collection year, Cronbach's coefficient alpha was computed for the maximum average rating samples within each setting and language. The results of these analyses for the English ratings, summarized in Table 11, show that all coefficients are .78 or greater, and support the high reliability of this rating index.

Table 9

Language Sample Tape Ratings:

Analysis of the Difference between the Minimum and Maximum Average Scale Ratings for Each Language and Each Setting Over Collection Years

Language	Setting	Percentage of Cases: Differences > 0.5	Percentage of Cases: Differences > 1.0	Total N
English	Classroom	12.1	1.9	481
	Playground	4.1	1.8	444
	Home	3.3	0.3	333
Spanish	Classroom	1.7	0.3	293
	Playground	2.7	0.7	439
	Home	0.6	0.0	338
	Average	4.1	0.8	388
	Total			2328

Table 10

Language Sample Tape Ratings - English:

Descriptive Statistics on Maximum Average Scale Ratings When Ratings from Two Settings Are Available

Setting Combination	Instructional Year	Average Rating: First Member Mean	First Member SD	Second Member Mean	Second Member SD	N
Classroom-Playground	0	4.8	.47	4.8	.42	54
	1	4.7	.49	4.7	.51	69
	2	4.6	.58	4.8	.42	86
	3	4.6	.49	4.8	.33	19
	4	4.6	.41	4.8	.31	12
Classroom-Home	0	4.7	.55	4.9	.32	42
	1	4.7	.44	4.7	.51	71
	2	4.7	.50	4.7	.38	63
	3	4.5	.56	4.5	.70	24
	4	4.8	.36	4.8	.22	5
Playground-Home	0	4.8	.43	4.9	.35	48
	1	4.7	.54	4.8	.43	70
	2	4.9	.31	4.7	.32	60
	3	4.8	.31	4.7	.65	21
	4	4.8	.31	4.8	.20	6
Total						650

Table 11

Language Sample Tape Ratings - English:

Reliability Analysis of the Total Scale Score for the Maximum Average Rating Samples for Each Setting for Each Data Collection Year

Setting	Collection Year	N of Cases	N of Items	Item Total Mean	Item Total SD	Standardized
Classroom	1	26	5	16.4	5.2	.96
	2	68	5	20.0	4.0	.94
	3	134	5	24.3	1.8	.89
	4	133	5	23.1	2.5	.90
	5	120	5	22.3	3.7	.97
Playground	1	8	5	21.5	2.3	.83
	2	50	5	21.8	4.3	.96
	3	145	5	24.4	1.3	.83
	4	127	5	24.1	1.3	.78
	5	114	5	24.4	1.4	.88
Home	1	12	5	19.7	4.4	.96
	2	41	5	22.2	2.9	.90
	3	102	5	24.3	1.5	.84
	4	98	5	24.1	1.5	.83
	5	80	5	23.5	2.3	.91

Table 12 presents the results of the reliability analysis of the Spanish ratings. These ratings show marked ceiling effects relative to the English ratings (larger means and smaller standard deviations coupled with many instances of deleted items due to no variance), and thus, show reduced reliability coefficients. However, in all cases where these effects are less severe, with the exception of the Year 1 Classroom ratings, the coefficients are quite acceptable.

Descriptive Statistics

As noted above, the taped interactions were taken in three communication settings: in the classroom, on the playground, and in the home. In all but the classroom setting, the student had relative freedom to communicate in the language of his choice. As can be seen in Table 13, the students tended to interact in the language(s) in which they had considerable skill. In those interactions in which English was used, the students' performance was rated native or near-native, as reflected by the 4+ ratings. Similarly, the interactions in Spanish reflected native speech (5.0). The growth rates are negligible since the students had mastered most of the structure and functions of the language to communicate successfully in the language of their choice.

The taped interactions allowed us to examine the patterns of language choice over time at each of the sites. Displayed in Table 14 are the descriptive findings relative to these patterns. The table is organized along the left margin by sites and by language (English, Spanish, Both). The choice of language is presented by setting for both entry (B) and exit (E) points of the study. The statistic reported is the percentage of children whose taped interaction reflected a particular language choice. For example, at Site 0, at the beginning of the study, 35% of the children's tapes were primarily in English, 14% primarily in Spanish, and 51% contained sufficient usage of both to rate the student's interaction in both Spanish and English. At exit (mostly 4th grade), however, 45% of the tapes in the classroom were primarily in English, none were primarily in Spanish, but 55% contained usage of both languages.

The general picture at Site 0 is one in which both English and Spanish are used in the classroom, with greater use of English occurring at the later grades. Spanish is the language of the playground, however, with a slight shift toward English at the end of the study. Spanish is maintained in the home, with a slightly greater use of English occurring in that setting toward the end of the study.

In Site 1, the pattern of language choice in the classroom is similar to that of Site 0; both languages are used by about one-half of the students, however, approximately 10% of the students were still using primarily Spanish in the classroom at the end of the study. On the playground, primarily English or use of both languages is found in the tapes of the majority of the students. Spanish is maintained in the home, with a slight shift to greater use of both in this setting at the end of the study.

Table 12

Language Sample Tape Ratings - Spanish:

Reliability Analysis of the Total Scale Score for the Maximum Average Rating Samples for Each Setting for Each Data Collection Year

Setting	Collection Year	N of Cases	N of Items*	Item Total Mean	Item Total SD	Standardized
Classroom	1	15	4	18.9	1.1	.45
	2	53	4	19.7	.8	.50
	3	76	2	9.9	.2	.00
	4	96	5	24.1	1.4	.78
	5	53	1	-	-	-
Playground	1	36	5	24.3	1.7	.80
	2	96	5	24.6	1.1	.77
	3	119	4	19.9	.4	.80
	4	116	5	24.6	.7	.57
	5	72	4	19.8	.6	.71
Home	1	28	5	24.0	1.7	.62
	2	89	5	24.6	.9	.69
	3	99	5	24.8	.8	.83
	4	73	5	24.4	1.1	.77
	5	49	4	19.4	1.1	.84

*Items with no variance were deleted from the analysis. For each deleted item, its mean was 5.0 (i.e., all respondents were rated at the top of the scale).

228

Table 13

Taped Interactions Language Samples:
Descriptive Statistics on Growth Measures from the Entire Sample and for Each Language Entry Group

Instrument	Measure	Statistic	Overall	Low Eng.	High Eng.	Low Span.	High Span.
LSTRE	S-Intrcp	M	4.4	4.3	4.5	4.7	4.3
LSTRE	S-Intrcp	S	1.1	1.3	0.9	0.9	1.1
LSTRE	S-Intrcp	N	153	59	94	60	93
LSTRE	Slope	M	0.1	.0	0.1	.0	0.1
LSTRE	Slope	S	0.3	0.5	0.2	0.3	0.4
LSTRE	Slope	N	153	59	94	60	93
LSTRS	S-Intrcp	M	5.0	5.0	5.0	5.0	5.0
LSTRS	S-Intrcp	S	0.2	0.2	0.3	0.3	0.2
LSTRS	S-Intrcp	N	119	64	55	23	96
LSTRS	Slope	M	.0	.0	-0.1	.0	.0
LSTRS	Slope	S	0.2	0.2	0.2	0.2	0.2
LSTRS	Slope	N	119	64	55	23	96

Table 14

Taped Interactions Language Samples:
Percentages of tapes by Language Choice for Each Site

Site	Language Use	Classroom Beginning	Classroom End	Playground Beginning	Playground End	Home Beginning	Home End
0	English	35	45	4	12	9	15
	Spanish	14	0	75	55	66	53
	Both	51	55	22	33	26	32
1	English	36	36	9	27	9	9
	Spanish	27	9	36	18	73	55
	Both	36	56	55	55	18	36
2	English	56	67	29	32	22	31
	Spanish	41	33	50	50	59	44
	Both	4	0	21	18	19	25
3	English	95	100	100	98	83	100
	Spanish	5	0	0	0	3	0
	Both	0	0	0	2	15	0
5	English	76	82	45	66	46	52
	Spanish	16	15	15	13	33	35
	Both	7	3	40	21	21	13

Site 0 (K-4): N=51,51,47
1 (K-4): N=11,11,11
2 (K-4): N=27,34,32
3 (K-2): N=42,42,40
5 (K-2): N=67,62,52

230

At Site 2, a larger number of students used primarily English in the classroom throughout the study than did the students in Sites 0 and 1. However, 33% of the students were still using primarily Spanish in the classroom interactions at the end of the study. Use of primarily Spanish or primarily English accounted for approximately 80% of the interactions on the playground. Spanish is the language of the home for the majority of the students, but a greater use of English is noticeable at the end of the study. At this site, the students tend to use primarily English or primarily Spanish in their interactions (as opposed to use of both languages) to a greater extend than do the students at Sites 0 and 1. This is particularly true in the school settings.

Site 3 differs considerable from the other four sites in the study. The language choice of the students at this site is almost exclusively English in all settings.

At Site 5, English is used extensively in all settings. However, in the home Spanish is maintained by about one third of the students, and both languages are used on the playground and in the homeby a sizeable number of students.

Summary

Adequate and accurate assessment of the oral language abilities of young students has long posed a challenge for practitioners and researchers alike. Objective measures, such as the currently-available standardized oral language proficiency tests, have been widely critized. The widespread dissatisfaction with these measures arises from the belief that these tests do not reflect the totality of the language resources that children possess, nor do they adequately predict children's ability to perform in the school setting. The inadequacy of such tests is undoubtedly due in part to the present state of knowledge, which at best is only partial or incomplete with respect to what constitutes language proficiency. Further dissatisfaction arises from the concern that formal testing of young children's language may in fact be measuring many things other than language (e.g., general readiness for school; knowledge of test-taking).

Subjective measures, such as teachers' ratings, have been maligned by some who point to the "human element" that comes into play with such procedures. Are criteria the same for each rater? How skillful are teachers in their ability to judge student performance in relation to the student's actual language resources? Are teachers influenced when making their ratings by how the student performs academically rather than by his language abilities per se?

Natural, or free speech, samples avoid some of the potential pitfalls of other types of measures, but they,too, have their limitations. How much speech should be collected, and how often should one collect such speech to be reasonably certain that the full range of a student's language abilities has been captured? Are those collected representative of the speech activities normally encountered by the student?

The SEDL research staff, fully aware of the limitations of the various kinds of measures and of the hazards involved in oral language assessment, employed multiple measures in an attempt to obtain a reasonablely accurate index of each student's oral language abilities and patterns of language choice over time. To the extent that we have been able to do this, a number of statements can be made:

- the students in the sample, on entry into school, varied considerably in the their degree of bilingualism

- the students, generally, made considerable in progress in acquiring skill in English; less growth was observable in their performance in Spanish

- site differences were apparent in the students' facility in Spanish and in English on entry and to their subsequent growth in each of the languages

- site differences were also observed in the patterns of language choice, both at entry and over time

- the language measures used provided information that illustrates the need for further research on effective means for assessing the oral language proficiency of young students. When compared to the teachers' ratings, the oral language proficiency test appeared to underestimate the students' ability, in both languages, on entry and to overestimate their ability at the higher grades.

Problems of language assessment notwithstanding, a number of factors may account for the growth patterns shown in the data. First, on entry into school, the students' new environment provides a wide exposure to English and to the formal aspects of language; thus, the opportunity as well as the necessity to learn English becomes greater, resulting perhaps in strong motivation to acquire English skills.

Secondly, instructional decisions made at the time of entry and thereafter are undoubtedly a contributing factor. All of the students in the bilingual sample were in bilingual programs at the time of student selection (kindergarten or first grade). However, some were mainstreamed to English medium classes after one year; others remained in bilingual classes throughout the years of the study.

Finally, the nature of the instructional program and of the reading program to which the students were assigned undoubtedly affected the student's growth, or lack thereof, in English and Spanish. Even though all of the students in the bilingual sample were assigned to bilingual programs at their entry into the study, not all of the students began initial reading instruction in Spanish; some received instruction in both languages concurrently, others were transferred to English reading instruction after one semester in first grade, others received their reading instruction entirely in English, and some were provided reading instruction primarily in Spanish from two to four years. Moreover, the character of the reading instruction differed from site to site and from classroom to classroom (see Volume VI: Instruction). For example, instruction in some classrooms focused heavily on letter-sound correspondence in the early stages of instruction; in others instruction was provided in the various components of reading from the early stages onward. Some classrooms were successful in promoting effectively language and literacy development. There is evidence in the data that it is certainly possible for bilingual children to thrive in schools and that by the time some of these children reach second and third grade they are proficient speakers of two languages and are fluent readers in both Spanish and English. However, it takes (among other things) skillful teaching and attention to the development of language in a variety of contexts within the school.

Other factors outside of the school most certainly played a role in shaping the growth patterns of the students. These include (among others) locale and the extent to which the two languages are used in the community, as well as the role of the home language in the affairs of the home and of the community; attitude of the student and others toward the maintenance of Spanish and/or the acquisition of each of the languages; and the extent to which written materials and formal usage are available to the students in each of the languages.

References

Burt, M. K., Dulay, H. C., & Hernández-Chávez, E. (1973). The bilingual syntax measure. New York, NY: Harcourt Brace Jovanovich, Inc.

De Avila, E. A., & Duncan, S. E. (1977). Language assessment scales - LAS I (2nd ed.). Corte Madera, CA: Linguametrics Group, Inc.

Duncan, S. E., and De Avila, E. A. (1981). Language assessment scales: Tests of oral language proficiency in English and Spanish (publicity flyer). San Rafael, CA: Linguametrics Group, Inc.

Mace-Matluck, B. J., Tunmer, W. E., & Domínguez, D. (1978). Oral language proficiency rating scale. Austin, TX: Southwest Educational Development Laboratory.

Mace-Matluck, B. J., Tunmer, W. E., & Domínguez, D. (1979). Ethnographic verification of language status. Austin, TX: Southwest Educational Development Laboratory.

Rivera, C., & Simich, C. (1982, March). Ethnographic/Sociolinguistic issues and the assessment of bilingual students' language proficiency. Paper presented at the meeting of the American Educational Research Association, New York, NY.

Texas Education Agency. (1978). State plan for bilingual education: A system for meeting the needs of limited English-speaking students in Texas.

Texas Education Agency. (1981). Comprehensive instruction, Title 19, Part II: Texas administrative code and statutory citations. Austin, TX: Author.

Woodcock, R. W. (1981). Batería Woodcock de proficiencia en el idioma - Versión en Español. New York, NY: Teaching Resources.

APPENDIX A

Description of the Bilingual Syntax Measure and the Batería Woodcock de proficiencia en el idioma - Versión en Español

THE BILINGUAL SYNTAX MEASURE

The Bilingual Syntax Measure - BSM was developed "to provide an instrument to measure oral language proficiency" which will allow inferences to be made "about a child's language dominance, the level of second language acquisition, and the degree of maintenance or loss of the first language" (Burt, Dulay, Hernández-Chávez, 1976, p. 13). The BSM exists in two levels in both English and Spanish. Level I is designed for use in Kindergarten through Grade 2. Level II was not used by any of the districts in the study, therefore, it will not be treated here.

Each test is administered individually. Administration time required is approximately 10 to 15 minutes per test. Trained school personnel, who were proficient in the language being tested, administered the BSM to one cohort of kindergarten students in one of the school districts in the study. This occurred as a normal procedure in the district's process of identifying limited English-speaking students for program placement. The materials consist of an administration manual in each language, the Picture Booklet containing the stimuli for the items, a response booklet for each child in the appropriate language (English or Spanish), and a technical manual.

The theoretical framework underlying the BSM is derived from the assumption that children acquire a second language by a process of "creative construction." Syntax was chosen as the measure of proficiency "because it is more stable across idiolects and dialects than vocabulary, pronunciation, or the functional uses of language. BSM items were constructed to elicit natural speech in English and in Spanish which could then be measured for syntactic proficiency" (Burt, Dulay, & Hernández-Chávez, 1976, p. 13). The test consists of 25 items.

In the first five items, the student is shown a 8 1/2" x 11" brightly-colored picture. The examiner points to a part of the picture and poses the stimulus question. The student's response is recorded verbatim for each item for which lines are provided (four of the five items). If the child has not responded to at least three of the test questions, the test is discontinued at that point. If the child has responded to at least three test questions, the examiner proceeds to the next set of items. For items #6 through #9, the child is shown simultaneously the previous picture and an additional picture which extends the scene presented earlier. The examiner again points to the appropriate section(s) of the pictures and poses the stimulus question and records the student's response for each of the four items. A new picture is presented for items #10 through #18, and the student's responses are recorded for those questions for which lines are provided (seven of the nine). For the final five items, three new pictures are presented and the test procedure described above is continued, with student responses recorded for three of the five items.

Following completion of the test, the scorer evaluates the student's response for each item recorded and, on the basis of criteria provided on the final page of the student's response booklet, assigns one of five levels: Level 1 - No English (Spanish); Level 2 - Receptive English (Spanish) only; Level 3 - Survival English (Spanish); Level 4 - Intermediate English (Spanish); Level 5 - Proficient English (Spanish).

BATERÍA WOODCOCK DE PROFICIENCIA EN EL IDIOMA - VERSIÓN EN ESPAÑOL
(The Woodcock Language Proficiency Battery - Spanish Form)

Woodcock Language Proficiency Battery - Spanish Form (WLPB-Span) is designed for use with Spanish-speaking students who have English as their second language. The instrument provides "an overview of their [the students'] Spanish language skills, which can aid instructional planning" (Woodcock, 1981b, p. 8). The battery exists in both English and Spanish, however, only the Spanish edition was used in the present study with the monolingual Spanish-speaking students in the Northern Mexico site.

The battery is designed for use with students ranging from preschool youngsters through university students and adults. In all subtests, the items are arranged in order of difficulty, with the easiest item first and the most difficult item last. The "operating range" is the set of consecutive items below which the student has essentially a 100% chance of getting all items correct (the basal level) and above which the student has virtually no chance of getting any items correct (the ceiling level). The goal of the testing is to start at an appropriate point within the student's operating range and then continue testing until all items within the operating range have been administered. The rules for obtaining basal and ceiling levels are included at the beginning of each subtest in the test book and are stated briefly at the top of each subtest in the Response Booklet. An example of the procedure used for determining these levels are described below as each of the subtests is discussed.

E   est is administered individually. Approximately 45 minutes is required to administer all eight subtests. Trained SEDL data collectors, who are native speakers of Spanish, administered the test battery to the students in the study at the Northern Mexico site. The materials consist of an examiner's manual; a test book which contains the test stimuli, both verbal and visual; and a Response Booklet for each student.

The WLPB-Span samples a wide range of language skills (oral, reading, written) normally required for performance in school. The test consists of eight subtests: Picture Vocabulary, Antonyms-Synonyms, Analogies, Letter-Word Identification, Word Attack, Passage Comprehension, Dictation, Proofing (Punctuation and Capitalization, Spelling, Usage).

The first subtest, Picture Vocabulary, requires the student to provide a verbal label for pictured objects or actions. There are 33 items in this subtest. For Preschool through Grade 3 students, the test is begun with Item #1 and proceeds until the student responds incorrectly to five consecutive items (ceiling level). Older students begin with more difficult items; for example, Grade 4 through Grade 6 students begin with Item #4, and adults begin with Item #15. Each item administered is scored by placing a "1" (correct response) or a "0" (incorrect or no response) in the appropriate space in the Response

Booklet. If a student provides an answer that cannot be scored either "1" or "0," probes are provided to assist in clarifying the student's response. The raw score for the subtest is the number of items correct and is obtained by counting as correct all items below the basal level (five consecutive items correct) plus all items scored as correct in the student's operating range (from basal level to ceiling level). Essentially the same scoring procedure is used for all of the subtests, with some variation required in the subtests of written language.

The second subtest, Antonyms-Synonyms, measures the student's knowledge of word meanings. Part A (Antonyms) requires the student to state a word whose meaning is the opposite of the stimulus word presented by the examiner. Part B (Synonyms) requires the student to state a word whose meaning is approximately the same as the presented word. The Antonyms portion of the test contains 32 items; the synonyms portion contains 25 items. There are no visual supports for this subtest.

In the third subtest, Analogies, the student is required to complete phrases with words that indicate appropriate analogies (e.g., Los gatos andan; los peces ..... nadan). Thirty-seven items are included in this subtest.

The fourth subtest, Letter-Word Identification, requires the student to identify (read) letters and words that appear in large type on the student's side of the test book. This subtest contains 50 items, the first four of which are letter-identification items; the remaining 46 items require the student to identify words.

In the fifth subtest, Word Attack, the student is required to read nonsense words (letter combinations that are not actual words). This subtest measures the student's ability to apply phonic and structural analysis skills in order to pronounce words that may be unfamiliar. The 28 items of this subtest were "selected so that almost all phonemes in the Spanish language are represented by at least one major spelling pattern" (Woodcock, 1981b, p. 5).

The sixth subtest, Passage Comprehension, utilizes the cloze procedure. It measures the student's ability to use contextual information in a short passage to supply a key word missing from the text. Examples of appropriate responses are provided. There are 28 items in this subtest.

The seventh subtest, Dictation, requires the student to respond in writing to a variety of instructions requiring knowledge of letter forms, spelling, punctuation, capitalization, and usage. Most of the items require the student to write a single word or abbreviation in response to the examiner's instructions. All items are presented in a manner similar to a traditional spelling test. This subtest contains 42 items.

The final subtest, Proofing, requires the student to identify mistakes in typewritten passages and to indicate how to correct each mistake. The student is informed that each typewritten passage contains one and only one error. Errors include incorrect punctuation, incorrect capitalization, inappropriate form of a word, and misspellings. This subtest contains 40 items.

The WLPB-Span is designed so that a combined score from the Dictation and Proofing subtests may be obtained on the punctuation and capitalization items, the spelling items, and the usage items. Provisions are made in the Response Booklet for the examiner to score across the two subtests in respect to these components and to plot results on the subtest profile. The two written language subtests (Dictation and Proofing) contain 23 items measuring punctuation and capitalization skills. A score for spelling may also be derived from the 34 items from the Dictation and Proofing subtests. Similarly, the score for the usage items is compiled on the bases of 25 items drawn from the two subtests.

Although subtests are the basic component of the WLPB-Span, clusters of subtests provide the primary basis for interpretation. Four cluster scores are derived from the raw scores. An oral language performance score is obtained from the student's performance on the Picture Vocabulary, Antonyms-Synonyms, and Analogies subtests. This measure of oral language is based on the rationale that "the abilities required to derive meaning and produce meaningful responses in the execution of certain cognitive tasks are prerequisites to understanding and producing oral language" (Woodcock, 1981b, p. 40). A reading score is derived by clustering the student's performance on the Letter-Word Identification, Word Attack, and Passage Comprehension subtests. These subtests represent "three of the most significant aspects of overall reading ability" (Woodcock, 1981b, p. 40). The cluster score for written language is obtained from the student's performance on the Dictation and Proofing subtests. This cluster measures written language skills in two contexts (i.e., supplying the correct form upon demand and detecting errors in previously written material). An overall score is a composite score based on the three clusters described above, and its "primary function is to provide a general index of overall functioning in the Spanish language" (Woodcock, 1981b, p. 41).

To obtain the cluster scores and overall score, percent correct indices are computed for each of the subtests which are then averaged based on the clustering described above.

References

Burt, M. K., Dulay, H. C., & Hernández-Chávez, E. (1973). The bilingual syntax measure. New York, NY: Harcourt Brace Jovanovich, Inc.

Burt, M. K., Dulay, H. C., & Hernández-Chávez, E. (1976). The bilingual syntax measure: Technical handbook. New York, NY: Harcourt Brace Jovanovich, Inc.

Woodcock, R. W. (1981a). Batería Woodcock de proficiencia en el idioma - Versión en Español. New York, NY: Teaching Resources.

Woodcock, R. W. (1981b). Woodcock language proficiency battery - Spanish form: Examiner's manual. New York, NY: Teaching Resources.

APPENDIX B

Sample Score Sheet

Language Assessment Scales

BEST COPY AVAILABLE

Student Score Sheet
for
LANGUAGE ASSESSMENT SCALES
LAS – Level I English)

Name ______ Sex ______ Date of Test ______
Date of Birth ______ Age ______
Teacher ______

I. MINIMAL SOUND PAIRS

1. ...
2. ...
3. very-berry
4. five-five
5. yellow-yellow
6. yet-Jess
7. hit-hit – 78, 79
8. hop-up – 29
9. spun-spun
10. especially-specially
11. pet-pat
12. back-pack
13. deep-dip – 87, 88, 89
14. meat-meat 88, 89
15. ...
16. rung-rings
17. ...
18. thin-tin
19. chain-chain
20. shop-chop
21. rice-rise
22. ten-tan – 66, 67
23. set-set
24. send-sent
25. mold-mold
26. ...
27. mop-mop
28. cold-gold
29. whether-weather – 99, 100
30. rain-ray

II. LEXICAL

Teacher Instructions: Tell the student you will be showing him/her some pictures. Then, point to each picture, and ask "What's this?"

31. table
32. train
33. dog
34. spoon
35. sofa (couch, etc.)
36. ...
37. ...
38. ...
39. ...
40. ...
41. chicken
42. bread
43. hammer
44. submarine
45. dinosaur
46. watermelon (melon)
47. candle
48. airplane
49. camel
50. cheese

... 11, 18, 22, 31, 32, 33, ... 70, 72, 86, 88, 89, 97, 103, 111, etc.

III. PHONEMES

Student Instructions: Are you ready? I want you to say exactly what you hear on the tape.

Examples: If you hear dog, you say dog. If you hear, it's raining, you say ______.

Item		Activities*
51. this 52. My father is further.	}	1, 2, 3
53. very 54. The rivers are moving.	}	12, 13, 14
55. Yes 56. The yard is yellow.	}	18, 19, 20
57. I am 58. The hat is hot.	}	23, 24
59. tuck 60. He hugged the bug.	}	29, 30, 31
61. bad 62. He sat on a mat.	}	36, 37
63. stop 64. The star can spin.	}	40, 41, 42
65. thing 66. Old Kathy is thin.	}	46, 48, 51
67. cheap 68. He chewed his chocolate.	}	54, 121, 122
69. peas 70. The boys were busy.	}	61, 63, 64
71. bed 72. Let the pet in.	}	66, 66, 69
73. toad 74. The food was good.	}	71, 73, 74
75. hill 76. He bit the chip.	}	77, 78
77. rib 78. The crab was in the tub.	}	82, 83, 84
79. beet 80. They need the feed.	}	87, 88, 89
81. bag 82. My gum is good.	}	91, 92, 94
83. white 84. There's white and wheat.	}	98, 99, 100
85. paint 86. The pig was in the park.	}	95, 124

IV. SENTENCE COMPREHENSION

Student Instructions: Listen to the tape, then point to the picture that shows what you heard.

87. The forks are held by both children.
88. The man is pushed by the woman.
89. The girl is not on the bike.
90. The boy does not hold a duck.
91. The woman feeds herself with a spoon.
92. The men and women are very unhappy.
93. The woman is riding the horse and the little girl is watching.
94. The fattest little boy is sitting.
95. The boy is eating the girl's food.
96. The cat jumped and the dog sat.

Any or all of the following LAS® Language Arts Supplement games and activities would be appropriate for enrichment of oral comprehension: 9, 15, 17, 34, 38, 43, 45, 48, 49, 52, 57, 63, 65, 69, 71, 74, 79, 90, 93, 104, 115, 116, 123, etc.

V. ORAL PRODUCTION: (Storytelling)

Scored on back of this sheet.

*If item is missed, these activities in the LAS® Language Arts Supplement (English) would be appropriate.

SCORING CALCULATIONS

Prs		Lex.		Phon.		Comp.		Prod.		Sub-Total			Total	Level
☐	+	☐	+	☐	+	☐	+	☐	=	☐	x 100	=	☐	☐
1		2		3		4		5		6			7	8

BEST COPY AVAILABLE

LAS® I — Student Score Sheet

Name ______________________________ Date of Test ________

V Production — Storytelling

Student Instructions: Now we're going to have a story. See these pictures? Aren't they neat? Well, these pictures tell a story and I'm going to play the story for you now. After the story is over I'd like you to tell me the story, as closely as you can remember it, exactly the way it was on the tape. So you'll have to listen very carefully. Let's see how much you can remember. Ready?

(After playing tape): Now that was a pretty good story, wasn't it. Can you tell me the story exactly the way you heard it?

Teacher Instructions: Arrange test book so 4 pictures can be seen simultaneously as student listens to tape. After hearing tape, ask student to retell story. BE SURE TO WRITE DOWN EVERY WORD OF STUDENT RESPONSE EXACTLY AS GIVEN. If student does not produce at least 50 words, try probe questions such as examples given below. Again, write down response exactly as spoken.

Probe questions to be used if necessary

1. What did he like to do? (pointing to silly old monster)
2. What did he do one summer day?
3. What [illegible] after he drank the [illegible]?
4. [illegible]
5. [illegible]
6. [illegible]
7. [illegible]
8. [illegible]
9. What did the little [illegible] give him?
10. What [illegible] the silly old monster never [illegible] to do again?

[illegible] 5, 18, [illegible] 1 5, [illegible]

[illegible]

3. Describing his or her family [illegible] in English to a monolingual [illegible]
4. Explaining to a teacher in English why s/he had been absent from class
5. Describing a science experiment in English. 5 4 3 2 1
6. How [illegible]

BEST COPY AVAILABLE

Student Score Sheet
for
LANGUAGE ASSESSMENT SCALES©
LAS® – Level I (Spanish)

Date of Test ______________

Name ______________________ Sex __________ Date of Birth __________ Age ______

District ______________ School ______________ Teacher ______________

Grade ______________ Home Language ______________ Ethnic Group ______________

Examiner ______________________ Test Language ______________

I SONIDOS MINIMOS EN PARES: (MINIMAL SOUND PAIRS)

Instrucciones al estudiante: Cuando oyes dos palabras en esta cinta, dime si suenan iguales o diferentes.

Ejemplo: 1) libro, puerta – diferentes
2) equipo, equipo – iguales

____ 1. todo-toro – 28,50,37*
____ 2. derecho-derecha – 1,9
____ 3. coro-corro – 49,23,27
____ 4. pisada-posada – 11,107,17
____ 5. come-como – 42,108,45
____ 6. boca-boca – 1,108,109
____ 7. secar-sacar – 106,110
____ 8. mitad-mitad – 12,13,14
____ 9. pasar-pisar – 1,11,8
____ 10. sabido-subido – 3,82,10
____ 11. cuñado-cuñada – 83,84
____ 12. pecar-picar – 43,11,14
____ 13. temor-tumor – 42,84,46
____ 14. cinco-cinco – 11,12,14
____ 15. paño-baño – 66,54,59
____ 16. tía-día – 29,72,76
____ 17. zona-cena – 107,105,45
____ 18. pesada-pesada – 42,43,47
____ 19. buscar-juzgar 87,90,93
____ 20. pasa-pasa – 1,2,3
____ 21. cana-caña – 117,19,20
____ 22. muro-moro – 84,82
____ 23. masa-masa – 109,5,6
____ 24. forro-ferro – 23,24,25
____ 25. pidiendo-pudiendo – 13,84,18
____ 26. goma-goma – 80,81,96
____ 27. lugar-ligar – 84,13,18
____ 28. casa-gasa – 94,95,77
____ 29. caso-casa – 106,1,9
____ 30. poco-pico – 106,12,17

II LEXICO: (LEXICAL)

Instrucciones al maestro: Dígale al estudiante que él/ella va a ver un dibujo, luego señale cada fotografía y pregúntele: ¿Qué es esto?

____ 31. mesa
____ 32. tren
____ 33. perro
____ 34. manzana
____ 35. sofá
____ 36. bicicleta
____ 37. elefante
____ 38. plátano (banana)
____ 39. cuchillo
____ 40. cohete
____ 41. pollo (gallo)
____ 42. pan
____ 43. martillo
____ 44. submarino
____ 45. dinosauro (dinosaurio)
____ 46. sandía
____ 47. candela (vela)
____ 48. avión
____ 49. camello
____ 50. queso

Any or all of the following games and activities from the LAS® Language Arts Supplement would be appropriate for vocabulary enrichment: 3, 4, 13, 14, 16, 21, 24, 26, 30, 34, 40, 41, 51, 52, 56, 57, 62, 67, 79, 80, 84, 88, 92, 96, 97, 115, 120, 121, 122, etc.

III FONEMAS: (PHONEMES)

Instrucciones al estudiante: ¿Estás listo/a? Repite exactamente lo que oyes en la cinta.

Por ejemplo, si oyes casa, vas a decir casa. Si oyes, buenos días, vas a decir ______________.

____ 51. abre
____ 52. Está en la mesa. } 1,5

____ 53. piso
____ 54. Irma tocó el hilo. } 11,12

____ 55. araña
____ 56. Las niñas lo enseñan. } 19,20

____ 57. tierra
____ 58. El perro hace ruido. } 22,23

____ 59. dime
____ 60. No pido nada malo. } 28,29

____ 61. jugo
____ 62. Es el jabón de Julio. } 38,39

____ 63. entra
____ 64. No me lo deja. } 42,43,47

____ 65. tarde
____ 66. Es el arte de María. } 49,50

____ 67. bosque
____ 68. A ver si viene. } 54,55

____ 69. payaso
____ 70. El topo juega en el pozo. } 66,67,68

____ 71. tabla
____ 72. El pato come tomates. } 70,71,73

____ 73. goma
____ 74. No juegues en el agua. } 77,78,79

____ 75. uvas
____ 76. Las nubes no vuelan. } 82,83

____ 77. buscar
____ 78. El amo hace lo mismo. } 86,87,90,91

____ 79. queso
____ 80. ¿Quién vió el parque? } 94,96,100

____ 81. lomas
____ 82. El lobo bailó aquí. } 101,102

____ 83. oveja
____ 84. Se ven dos ojos. } 104,106,108

____ 85. llanta
____ 86. El camello me llamó. } 112,113

IV COMPRENSION DE FRASES (Sentence Comprehension)

Instrucciones al estudiante: Escucha la cinta, luego enséñame el dibujo que indica lo que oíste.

____ 87. El joven corre muy rápido.
____ 88. La niña está columpiándose.
____ 89. La mujer alta y delgada está sentada.
____ 90. El niño agarra el juguete de la niña.
____ 91. La señorita no tiene dinero.
____ 92. El hombre estaba comiendo cuando el árbol se cayó.
____ 93. El gato se sentó y el perro brincó.
____ 94. La niña pequeña ha roto el vaso.
____ 95. Todas las mujeres están muy contentas.
____ 96. La puerta es cerrada por los dos niños.

Any or all of the following games and activities from the LAS® Language Arts Supplent would be appropriate for enrichment of oral comprehension: 6,15,16,21,26,31,35,44,48,57,58,67,73, 80,84,86,89,109,112,114,118,119,123, etc.

V PRODUCCION ORAL – CUENTOS (Production – Storytelling)

Scored on back of this sheet.

*If item is missed, these activities in the LAS® Language Arts Supplement (Spanish) would be appropriate.

SCORING CALCULATIONS

Prs.		Lex.		Phon.		Comp.		Prod.		Sub-Total			Total	Level
[ ]	+	[ ]	+	[ ]	+	[ ]	+	[ ]	=	[ ]	x 100	=	[ ]	[ ]
1		2		3		4		5		6			7	8

LAS® I – Student Score Sheet

Name ________________________________ Date of Test __________

V Production – Storytelling

Instrucciones al estudiante: Ahora vamos a escuchar una historia. ¿Ves estos dibujos? Pues, los dibujos cuentan una historia y ahora puedes escucharla. Después de escuchar la historia, quiero que me repitas la historia, tan exacto como puedes recordarla. Así que tendrás que escuchar con mucho cuidado. Vamos a ver cuánto puedes recordar. ¿Listo(a)?

(After playing tape): Muy bien. ¿Me puedes repetir (contar) la historia exactamente como la oíste?

Teacher Instructions: Arrange test book so 4 pictures can be seen simultaneously as student listens to tape. After hearing tape, ask student to retell story. BE SURE TO WRITE DOWN EVERY WORD OF STUDENT RESPONSE EXACTLY AS GIVEN. If student does not produce at least 50 words, try probe questions such as examples given below. Again, write down response exactly as spoken.

Probe questions to be used if necessary:

1. ¿Qué le gustaba hacer? (pointing to giant) ________
2. ¿Qué hizo la giganta un día de vera ? ________
3. ¿Qué dijo la giganta después de comer la pintura? ________
4. ¿Quiénes son ellos? (pointing to giant's friends) ________
5. ¿Qué le preguntaron sus amigos? ________
6. ¿Qué le trajeron sus amigos? ________
7. ¿Qué le dio el gigante grande? ________
8. ¿Qué le dio el gigante mediano? ________
9. ¿Qué le dio la giganta pequeña? ________
10. ¿Qué es lo que giganta nunca más va a hacer? ________

Any or all of the following LAS® Language Arts Supplement games and activities would be appropriate for enrichment of syntax production: 16, 21, 24, 26, 31, 35, 41, 52, 58, 65, 69, 72, 74, 98, 105, 111, 118, etc.

OBSERVATIONS

Based on your observations, please give your assessment of this student's use of the Spanish language. Rate the student's probability of success in the following situations.

	(Succeed fully) 5–4–3–2–1 (fail) (Circle One)
1. Asking for directions in Spanish to an unfamiliar part of the school.	5 – 4 – 3 – 2 – 1
2. Telling a joke in Spanish to monolingual peers.	5 – 4 – 3 – 2 – 1
3. Describing his/her family composition in Spanish to a monolingual peer or teacher.	5 – 4 – 3 – 2 – 1
4. Explaining to a teacher in Spanish why s/he had been absent from class.	5 – 4 – 3 – 2 – 1
5. Describing a science experiment in Spanish.	5 – 4 – 3 – 2 – 1

6. How long have you known this student? ______________

______ Total ______ Average

APPENDIX C

Oral Language Proficiency Rating Scale

© SEDL ORAL LANGUAGE PROFICIENCY RATING SCALE

Betty J. Mace-Matluck
William E. Tunmer
Domingo Domínguez

With assistance from
Aaron Bar-Adon
The University of Texas at Austin

Division of Bilingual and International Education
SOUTHWEST EDUCATIONAL DEVELOPMENT LABORATORY
211 East 7th Street
Austin, Texas 78701

ORAL PROFICIENCY RATING SCALE

Criteria

PRONUNCIATION

1. Often unintelligible due to excessive mispronunciation, making comprehension extremely difficult.

2. Intelligible, but with frequent mispronunciations which may, at times, interfere with communication.

3. Always intelligible, but reflects occasional mispronunciations which are usually systematic.

4. Essentially like that of a native speaker, except for some residue or overtones that suggest non-nativeness.

5. For all practical purposes, like that of a native speaker; pronunciation may reflect characteristic features of the dialect of the region.

GRAMMAR

1. Makes excessive number of errors in grammar, except in stock phrases; extremely limited in range and variety of syntactic structures.

2. Makes frequent errors in grammar, which may interfere with normal communication; rather limited in range and variety of syntactic structures; frequently resorts to rephrasing in midcourse.

3. Makes occasional errors in grammar which may, at times, obscure meaning; range and variety of syntactic structures are relatively limited when compared with those of native peers.

4. Makes sporadic errors in grammar that are non-typical of native speakers of the same age; grammar is essentially like that of native speakers with syntactic structures resembling those of native peers in range and variety.

VOCABULARY

1. Vocabulary is severely limited and often hampers communication.

2. Vocabulary is limited when compared with native peers; frequent use of inappropriate terms.

3. Vocabulary is mostly adequate, but occasionally deficient.

4. Vocabulary is essentially like that of a native speaker of the same age, except for sporadic groping for appropriate terms.

5. For all practical purposes, vocabulary is like that of a native speaker of the same age.

COMPREHENSION

1. Understands very little speech, except for a limited number of items frequently used in the classroom or social setting (e.g., greetings); requires simplification, repetition, and/or much use of gestures.

2. Understands some adult or peer speech spoken at a normal rate, but often requires simplification of speech or frequent repetition or rephrasing.

3. Understands most adult or peer-group speech, spoken at a normal rate, that would usually be understood by native peers, but occasionally demonstrates lack of, or only partial, understanding.

4. Understands essentially everything, spoken at a normal rate, in school-related, social, or peer-group conversation, except for certain idiomatic phrases or conventionalized usage of the language.

5. Understands everything in both classroom and playgroup speech which would usually be expected of native speakers of the same age.

OVERALL COMMUNICATION SKILL

1. Is able to participate only minimally in school-related or peer-group conversations conducted in the language. Speech is generally characterized by labored production, incomplete sentences, and/or excessive number of errors.

2. Is able to get the gist of most school-related and peer-group conversations, but is unable to participate with facility in any but very familiar, routine conversations. Speech is frequently uneven, hesitant, and fragmented.

3. Understands and speaks the language adequately to participate in most school-related and peer-group conversations. Speech is characterized by occasional errors in grammar, some groping for words, and at times, hesitancy and unevenness in production.

4. Uses the language fluently and accurately, for the most part, and is able to participate successfully in all school-related and peer-group conversations. Speech, while smooth, effortless, and generally without error, contains some sound qualities and grammatical structures which suggest non-nativeness.

5. For all practical purposes, uses the language like a native speaker of the same age. Speech in all school-related and play-group conversations is smooth, effortless, and native-like in accuracy.

ORAL PROFICIENCY RATING SCALE

ENGLISH

Student's Name ______________________ Grade ____________

Teacher ______________________ Date ____________

School ______________________ Rater ____________

District ______________________________________

INSTRUCTIONS: Please refer to the accompanying criteria sheet and circle below the number corresponding to the statement which most accurately describes the student's level of proficiency for each of the language components indicated.

PRONUNCIATION	GRAMMAR	VOCABULARY	COMPREHENSION	OVERALL COMMUNICATIVE SKILL
1	1	1	1	2
2	2	2	2	2
3	3	3	3	3
4	4	4	4	4
5	5	5	5	5

252

ORAL PROFICIENCY RATING SCALE

SPANISH

Student's Name ______________________ Grade ____________

Teacher ______________________ Date ____________

School ______________________ Rater ____________

District ______________________

INSTRUCTIONS: Please refer to the accompanying criteria sheet and circle below the number corresponding to the statement which most accurately describes the student's level of proficiency for each of the language components indicated.

PRONUNCIATION	GRAMMAR	VOCABULARY	COMPREHENSION	OVERALL COMMUNICATIVE SKILL
1	1	1	1	2
2	2	2	2	2
3	3	3	3	3
4	4	4	4	4
5	5	5	5	5

253

APPENDIX D

Sample Rating Form

Taped Interactions - Language Samples

Inventory No. ____________

LANGUAGE SAMPLE RATING SHEET

Student's Name ______________________ Grade ____________

Teacher ______________________ Date Collected ____________

School ______________________ Rater ____________

District ______________________ Date Rated ____________

1. Type of interaction (circle one):[1] T P F
2. General language use of interlocuters (circle one):[2] S E A C B
3. General language use of student (circle one):[2] S E A C B
4. Oral proficiency rating:[3]

	SPANISH (if used by student)	ENGLISH (if used by student)
PRONUNCIATION	1 2 3 4 5	1 2 3 4 5
GRAMMAR	1 2 3 4 5	1 2 3 4 5
VOCABULARY	1 2 3 4 5	1 2 3 4 5
COMPREHENSION	1 2 3 4 5	1 2 3 4 5
OVERALL COMMUNICATIVE SKILL	1 2 3 4 5	1 2 3 4 5

1 T = Teacher-Pupil; P = Peer-Pupil; F = Family-Pupil

2 S = Spanish; E = English; A = Alternate use of both; C = Code Switching; B = Both

3 Refer to accompanying criteria sheet

Final Report

TEACHING READING TO BILINGUAL CHILDREN STUDY

Volume 5
Reading Growth

Wesley A. Hoover
Robert C. Calfee
Betty J. Mace-Matluck

Document BRS-84-R. 1-V

Preston C. Kronkosky
Executive Director

Southwest Educational Development Laboratory
211 East Seventh Street
Austin, Texas
(512) 476-6861

November 1984

Final Report

TEACHING READING TO BILINGUAL CHILDREN STUDY

Volume 5
Reading Growth

Wesley A. Hoover, Robert C. Calfee
Betty J. Mace-Matluck

Document BRS-84-R. 1-V

Preston C. Kronkosky, Executive Director
Southwest Educational Development Laboratory
Austin, Texas

November 1984

There were many individuals and institutions who contributed to this research effort. We wish to express our sincere gratitude to the parents, students, and school personnel who provided the necessary data from which this study is derived.

In addition, several other individuals made valuable contributions to the study, for which we are indebted: Robert C. Calfee, Sylvia C. Peña, and Blanca de Alvarez.

And finally, we wish to thank the local data collectors at the school sites, many of whom remained with the study throughout its duration: Ramiro Barrera, Beatrice Cantú, Irene Cavazos, Carolyn Cruz, María de Obregón, Gloria de Torres, Gigi Galván, Olga Hernández, Irene Méndez, Guadalupe Treviño, Rosalinda Villanpando, and Gloria Villarreal. Their patience, dedication, and hard work helped make this study a reality.

Betty J. Mace-Matluck
Wesley A. Hoover

TABLE OF CONTENTS

LIST OF TABLES

280

LIST OF FIGURES

266

PREFACE

In June 1978 the National Institute of Education (NIE) funded the Southwest Educational Development Laboratory (SEDL) to conduct a longitudinal study on the Teaching of Reading to Bilingual Children. Educators and policymakers alike have long recognized that the ability to read is essential for success in school, in work, and in life; yet many children from second-language backgrounds have trouble learning to read in schools today. The majority of these youngsters are from Spanish-language backgrounds and from low income families. Special programs designed to meet the needs of these children are provided in schools, but there is limited research evidence to guide the development, evaluation, and implementation of these programs. This study is intended to provide information that will result in greater insights into what constitutes a favorable learning environment for children from Spanish-language backgrounds, what instructional sequences and events promote successful and efficient learning of literacy skills, and what the language and literacy outcomes of current schooling practices are for a large sample of these youngsters.

The study was conducted during the years of 1978 through 1984. It is a comprehensive longitudinal investigation of the development of reading skills from kindergarten through fourth grade for a representative sample of more than 350 children from bilingual backgrounds, and for smaller samples of children who, on entry into school, were monolingual in English or Spanish. In this "natural variation" study, teaching and learning were carefully documented in field settings at the several sites.

The goals of the study were to (a) describe variations in both English and Spanish language ability of students living in bilingual communities, (b) document prevailing practices in reading instruction for bilingual students, and c) investigate the relations between the instructional program and student achievement for students with differing entry profiles.

Description of the Study

Surveys of the general and school populations r   il an increase in the number of students whose language resources a   not an ideal match to the language of the school. An important question for educational practice and policy centers around the school's responsibilities in this situation. Bilingual programs, English-as-a-Second-Language classes, classroom aides, and "sink-or-swim" approaches can all be found in practice today. From limited evidence now available, none of these techniques has emerged as the one best system.

Hispanics make up the largest and fastest growing school-age population today. The demographics for some states show that over the next decade they may constitute as much as a third to a half of the population. In the state of Texas at present approximately one third of the school children are from Hispanic backgrounds (approaching one

million). They are found in virtually ever school district in the state. Many of the school districts in the southern portion of the state serve school populations of which 75% to 99% of the children are from Spanish-speaking backgrounds and, on entry into school, are often limited in their ability to speak English and to profit from instruction in that language. This population is not restricted to the border areas, however. Large urban centers in the state report as much as 20% of their school population from Hispanic backgrounds, with a concentration of some 80% to 90% in certain of their schools.

It is well documented that, in general, children from Spanish-speaking backgrounds, for whatever reason, often encounter difficulty in our nation's schools; they do more poorly on standardized tests than does the general school population, and their dropout rate is high. Bilingual education, in which students are given instruction partially through the home language until they have attained sufficient proficiency in English to benefit from English-medium instruction, has been the principal approach recommended by the Office for Civil Rights to ensure access to equal educational opportunity for these children. Although many individual programs have had considerable success in improving the academic performance of language-minority students, it has not been demonstrated that these programs generally are reducing inequality of educational opportunity on the large scale that was envisioned.

Growth in reading comes about for most youngsters through formal classroom instruction. Understanding the development of reading, and knowledge of the critical variables that determine success or failure, depends on careful examination of the instructional program -- not just the label over the classroom door, but the program as actually implemented by the classroom teacher.

Educators have raised several issues about the most effective way to help bilingual children become proficient readers of English. These include (a) valid assessment of the student's ability in the languages of the home and of the school, (b) the optimal balance of formal instruction in both languages, (c) the most effective transfer from one language to the other, and (d) bilingual support within the classroom environment. A major thesis of the Teaching Reading to Bilingual Children study is that addressing these issues (and others) requires a comprehensive and ecologically-valid investigation of the linkage between the child's language and the language of instruction.

Design of the Study

To achieve the objectives of the study, considerable attention was given to the selection of schools, teachers and students, to the instruments for assessing language and reading achievement, and to the methods for evaluating the classroom instruction. Each of these topics is discussed briefly below.

Schools, Classes and Teachers

Twenty schools and 200 teachers from six school districts participated in the study. Included are variations in the nature of the reading program (a range from phonics-oriented to meaning-based), classroom organization (some self-contained, others team-taught), and grade structure (the range of grades in the individual school and the extent of cross-grading both vary). The schools differed in size, SES, urbanicity, locale, and makeup of the student body (from medium to high concentration of bilingual students).

Student Cohorts

The study was undertaken in four cohorts or "waves" of students. Three of the cohorts consisted entirely, or in large part, of bilingual students. The first cohort was small (N=40) and of limited generality; the second was somewhat larger (N=80) and covered a slightly broader array of contexts. The third cohort which was both larger (N=200) and broader in its generality, incorporated a number of procedural improvements based on previous experience in the study and included a monolingual English-speaking sample. The fourth cohort consisted of a relatively small sample (N=60) of monolingual Spanish-speaking students.

All of the bilingual sites were from the state of Texas, as were the monolingual English-speaking students. The monolingual Spanish-speaking students were from one site in Northern Mexico.

The original design of the study called for each student to be assessed and observed from entry to kindergarten through exit from third grade. By covering the full range of the primary years, we would be able to examine the transition from "learning to read" through "reading to learn." For students in programs where the initial stages of reading were in Spanish, we also considered it important to determine the transition to competence in English reading.

The original design was in fact implemented for the first two cohorts; some of the students were tracked from first through fourth grade, but most followed the intended design. Due to limited funding in the later stages of the study the last two cohorts could not be followed for the full four years that were originally intended. The bilingual and monolingual English samples from the Texas sites were observed from kindergarten through second grade, and the monolingual Spanish samples from the site in Northern Mexico were observed from first through third grade (the program did not provide a kindergarten).

The monolingual samples were incorporated in the design to aid in validating the instruments for student assessment. Both the English and Spanish cohorts are small and not selected to be fully representative of monolingual populations. Data from these samples will be presented in Volume 3, as part of the discussion on the adequacy of the instruments for measuring growth. The study was designed to study the course of reading in bilingual students, not as a basis for comparing these students with monolingual youngsters. Accordingly, comparisons

between the various samples will not be made in this report, nor do we recommend that others attempt such comparisons.

Language Assessment

Several types of data were collected for each student on English and Spanish proficiency. Each year, early in the Fall and again in the Winter and Spring, teachers rated their students' language skills. Oral language proficiency tests were administered in the Fall of each year. Finally, audiotaped speech samples were obtained monthly on a rotating schedule in three settings: in the classroom, on the playground, and in the home.

Reading Assessment

Several instruments were used to measure reading achievement. Standardized test scores (mostly English) were collected yearly. More detailed information was obtained from a battery of individually-administered "performance based tests" in both English and Spanish. In kindergarten, the Stanford Foundation Skills Test was employed to measure the child's pre-reading skills. From the end of first grade on, the Interactive Reading Assessment System was administered during the Spring of each school year. This instrument provides independent measures of the student's skills in decoding, word meaning, fluency in oral reading, and comprehension. Finally, informal reading inventories were administered throughout the school year.

Classroom Observations and Teacher Interviews

Project staff conducted monthly observations of the reading instruction in each classroom and interviewed the teachers quarterly about their instructional plans. The observation instrument documented staffing patterns, grouping and organization, time allocation, the language of instruction, the character of instruction, the materials and procedures used, and the response of the students. The interviews focused on the teacher's general instructional objectives, as well as the objectives for individual target students. Taken together, these two instruments yield a rich characterization of the classroom environment for the target students.

Student Entry Variables, Classroom Factors, and Reading Achievement

The primary goals of the analyses were to identify the general relationships that characterize variation in these factors and to look for underlying regularities that are associated with success and failure, both in the early stage of reading instruction and in the year-to-year variations.

Documents

This report is one of a series of eight documents contained in the Final Report submitted to the National Institute of Education. A com-

plete list of these documents is provided on the inside of the cover of this report.

The study was a collaborative effort among a number of individuals and institutions. All members of the research team contributed to the thinking, planning, and writing of this series of documents, however, the individual whose name appears first in the list of authors was responsible for preparing the particular document.

Betty J. Mace-Matluck
Wesley A. Hoover
Co-Principal Investigators

Austin, Texas
November 30, 1984

INTRODUCTION

A primary purpose of the study was the investigation of patterns of growth in reading achievement. The study employed multiple measures for assessing each of the major components of skilled reading (vocabulary knowledge, decoding, and text comprehension), and for the bilingual sample, monitored such growth in both English and Spanish. In the early grades, the Stanford Foundation Skills Test was administered to assess components of "reading readiness"; in later grades, the Interactive Reading Assessment System was used to measure the components of skilled reading. Two additional indices of literacy were also employed: standardized reading achievement scores were collected yearly wherever available, and monthly progress in reading was monitored through an Informal Reading Inventory. In this volume, these instruments, and the data obtained from them, are discussed.

READING GROWTH

The pre-reading and reading instruments are discussed below, providing details of the tasks, materials, scoring, reliability, and descriptive statistics for the bilingual sample's performance. First, the pre-reading measures are treated, then those of reading, and finally, the relations between the two. Student performance on the informal reading inventories is not treated in this report; however, a description of the instrument is included in Appendix A.

Pre-reading Measures

Stanford Foundation Skills Test

The Stanford Foundation Skills Test - SFST (Calfee & Associates, 1978) is designed to measure the set of perceptual and language skills that provide the foundation for the acquisition of reading. The test, which is individually administered to pre-reading students in kindergarten and early first grade, includes subtests of alphabet knowledge, word naming, visual matching, auditory-phonetic segmentation, vocabulary, and story comprehension. Over ten years of development, the instrument has been expanded to include word definition and story comprehension tasks (Calfee & Associates, 1980) in order to improve its match with the companion Interactive Reading Assessment System (discussed below), which is appropriate for assessing developing reading skills.

Each of the two major revisions of the SFST have been converted into parallel Spanish-language versions (Calfee, & Pena, 1978; 1980). Within the two language versions, the non-linguistic materials employed (e.g., visual matching) are identical. Linguistic materials, however, are not direct translations between the versions (unless otherwise noted in the task descriptions below), but rather were generated employing the same procedures used to select the linguistic materials for the original versions of the test in English.

The only revision of the SFST during the data collection phase of the study was made after the administration given in Year 2, but prior to Year 3 testing. Thus, given the cohort structure discussed earlier, target students from Sites 0, 1, and 2 were tested with the first version, and targets from Sites 3, 4, and 5 were tested with the second version.

For the pilot targets selected in Collection Year 1, each student was administered only one language version of the SFST, left to each student's choice. For the remaining targets in the bilingual sample, all students were tested with both language versions. For these students, order of test administration was counterbalanced within each site, with approximately two weeks between testing. All students were tested in the Fall (October to November) of their kindergarten year, except for those bilingual targets from Sites 0, 1, and 2, and the monolingual-Spanish targets from Site 4, who were initially selected as first-graders, and tested in the Fall of first grade. Table 1 summarizes the language administrations for the 333 target students by site and grade level.

The specific structure for each of the SFST subtests, as well as the scoring procedures employed, are presented below. The order of the task descriptions follows the order in which each of the tasks was administered. These descriptions are followed by discussions of the SFST reliability assessments, descriptive statistics for the target sample's performance, and both intra- and inter-test correlations.

Tasks, Materials, and Scoring

The SFST consists of six major components: alphabet knowledge (both production and recognition), word naming, letter matching (both single and double letter items), auditory-phonetic segmentation, vocabulary (both fine distinctions and common labels), and story comprehension (listening only). A description of each of these tasks, including any revisions made, their scoring, and the derivation of summary measures, are discussed individually below.

Alphabet Knowledge. A young student's knowledge of letter names has been established as an important predictor of later success in reading, though the meaning of this relation is far from clear (Venezky, 1975). The studies of this phenomenon are almost exclusively from English language situations, and we know of no studies investigating this relation in the Spanish language.

The SFST includes two indices of Alphabet Knowledge. In the first task, the student is shown each of the capital letters (26 in the English version and 30 in the Spanish version), and is asked to name them. If the student makes no attempt to name six letters in a row, or, having been presented with each of the letters, fails to give the correct name for half of them, then a recognition test is administered. Here, the examiner says each letter name, and the student is asked to point to the corresponding letter on a response sheet containing all of the letters. Again, if the student fails to respond to six letters in a row, the testing is discontinued.

273

Table 1

Stanford Foundation Skills Test:

Distribution of Test Administrations for Target Students by Language Group, Site, Grade, and Language of Test

Language Group	Site	Target Students: Kindergarten English Only	Kindergarten Spanish Only	Kindergarten English & Spanish	First Grade English Only	First Grade Spanish Only	First Grade English & Spanish	Missing Data	Total
Bilingual	0	1	16	15	4	16	1	2	55
	1	-	-	7	-	-	9	1	17
	2	-	1	16	-	-	19	-	36
	3	-	2	67	-	-	-	7	76
	5	-	1	56	-	-	-	13	70
	Total	1	20	161	4	16	29	23	254
Monolingual English	3	9	-	-	-	-	-	-	9
	5	16	-	-	-	-	-	11	27
	Total	25						11	36
Monolingual Spanish	4	-	-	-	-	43	-	-	43

In the first version of the test, letters in the naming task were presented in a pre-determined random order (for a full description of the randomization procedure for the English alphabet, see Calfee, 1970; the Spanish alphabet was ordered in a similar fashion); the letters in the identification task were displayed in the same random order, but were named by the tester in a different pre-determined random order. In the second version of the test, letters in the naming task were presented by order of difficulty from easiest to most difficult (based on rankings obtained from the English and Spanish tests administered during the first two years); in the identification task, letters were arranged in the same random order as used in the previous version, but were named by the tester in the difficulty ordering used in the naming task.

For both the naming and identification tasks, each item was scored as correct ("C"), incorrect ("W"), no response ("N"), or assumed failure ("F"), the latter for letters not presented after six consecutive failures to respond. In deriving summary scores, correct responses ("C") were assigned a value of 1, and incorrect responses ("W", "N", and "F"), a value of 0. Performance on each task was summarized by the percent correct. For students who met the criterion of correctly naming at least half the letters in the naming task, and thus, were not tested on the identification task, a value of 100% was assumed on the latter. Overall performance on Alphabet Knowledge was the average percent correct over the two tasks.

Word Naming. Materials for this task consisted of 12 monosyllabic, familiar words selected from lists provided by experienced kindergarten teachers who were asked to give 10 words which their students were most often able to read. Upon completion of the alphabet tasks, each student was shown each of the 12 words, one at a time, and asked to read them aloud for the tester. If the student made no response to three words in succession, testing on the task was stopped. This task was added to the SFST for use in the Year 3 data collection phase, and thus, only targets in Sites 3, 4, and 5 were tested with it.

The scoring of this task matched that used in Alphabet Knowledge, employing the codes of correct ("C"), incorrect ("W"), no respons ("N"), and assumed failure ("F"). A value of 1 was assigned to items given correctly ("C"), and a value of 0 for all other items ("W", "N", and "F"). The percent correct over the 12 items served as a summary index of performance.

Letter Matching. Much has been written about the importance of visual perceptual skills in beginning reading (Gibson & Levin, 1975, but cf. Vellutino, 1980), and the SFST includes subtests that measure students' ability to match single letters and letter pairs. The materials are made from synthetic letters designed to incorporate the significant features of the Roman alphabet, while not appearing exactly like any particular letter (Calfee, 1977).

The Single-Letter and Double-Letter tests each consist of 12 items, where the student is shown a "target" in the center of a page, and is asked to mark any of the five alternatives (six in Double-Letter matching) that are an exact match to the target. The alternatives are presented in a semicircle around the target so that each is equidistant from it, and each item has at least one exact match. In the Single-Letter test, seven of the 12 items include a second exact match, as a way of assessing the student's care in examining all of the alternatives. Both of the tests include reversals: there are four items with "b-d"-like alternatives on the Single-Letter test, and all 12 items on the Double-Letter test have "was-saw"-like alternatives. The English and Spanish versions differ only in the language of the instruction set. No modifications to this subtest were made over the three collection years in which it was used.

Each item in these two tasks is scored for four response types, with 1 indicating a correct decision and 0, an incorrect decision: (a) whether or not a correct alternative was selected, (b) for items containing two correct alternatives, whether or not the second alternative was also selected, (c) whether or not a reversal alternative was selected, and (d) whether or not one or more non-reversal, but incorrect, alternatives were selected. Thus, for the 12 Single-Letter matching trials, a total of 35 response decisions were scored (i.e., 12 correct alternatives, seven second correct alternatives, four reversal alternatives, and 12 incorrect, non-reversal alternatives). The percentage of correct decisions (i.e., selection of correct alternatives and rejection of incorrect alternatives over these 35 items served as the summary index for this subtest. For Double-Letter Matching, 36 response decisions were scored (i.e., 12 correct alternatives, 12 reversal alternatives, and 12 incorrect, non-reversal alternatives), and the percentage of correct decisions also served as the summary index of performance.

Auditory-Phonetic Segmentation. Learning to read is partly a matter of learning to decode; that is, learning the relations between spelling patterns and phonetic (or phonemic) patterns. Numerous studies have documented the difficulties that youngsters experience in acquiring the concept of the "sounds" in a spoken word (e.g., Ehri, 1979), and the Auditory-Phonetic Segmentation test provides several indices of students' ability in this area

Because so few pre-readers have been explicitly exposed to the concept of phonetic segments in words, the test begins with a training sequence, after which the student is tested for transfer of the concept. The task is to delete the initial consonant segment from a presented word, and say aloud the vowel-consonant segment that remains. For instance, to the word "mice" the proper response is "ice"; to the word "rope", the correct answer is "ope".

The student is first shown a response card with pictures representing the vowel-consonant segments for the training series, and is familiarized with the names for each picture (eyes, ache, and eat). Training is carried out in a paired-associate fashion using the anticipation method; after a preliminary study trial, in which the student is shown nine stimulus words and told the correct answer, the same stimuli are then presented in random order, and the student is asked to try to anticipate the correct answer. The stimulus words are all consonant-vowel-consonant combinations, three paired with each of the responses. (The Spanish version employs words with a somewhat more complex syllabic structure because mono-syllabic words are fairly rare, but the segmentation principle is the same.) If the student's response to a given stimulus is correct, the tester reinforces the answer; if the answer is wrong, the tester gives the correct answer. Training continues until the student is correct on seven of the nine stimuli, up to a maximum of four trials through the entire nine-item list.

If the student successfully attains the criterion within the four training trials, a series of six transfer lists is administered. The first transfer list (T1) contains 15 stimuli -- the nine training stimuli plus six new stimuli that use the same responses as in training. The second list (T2) contains six items, employing six new stimuli with three new responses. For T1 and T2, the student has pictures available for response support, but for the remaining transfer trials, no picture supports for response are employed. For the third set of transfer trials (containing six items), a new set of picture stimuli are used, but the correct responses are all from the previous training and transfer episodes. T4, also containing six items, introduces all new stimulus words for which the previously learned responses are still used; the test is entirely oral from this time on (i.e., no picture supports for either stimuli or responses). On T5 and T6, each containing 12 items, the student is tested on new stimulus words, half of which are real and half of which are synthetic. The responses are novel, and consist of real words on T5, synthetic words on T6.

Each of the items in both the training and transfer segments of the subtest were scored as follows: (a) correct verbal response ("C"), (b) correct response by pointing to the correct picture, but without verbal response ("P"), (c) incorrect response, but one which rhymes with the stimulus ("R"), (d) incorrect response which is either a meaningful associate of the stimulus word or the stimulus word itself ("M"), (e) wild response ("W"), for oral responses not appropriately related to the stimulus, (f) no response ("N"), (g) assumed success ("S"), for items not tested in the training segment because a student meet the training criterion on an earlier trial, and (h) assumed failure ("F"), for items not tested in the transfer segments due to failure to meet the training segment criterion.

In deriving a summary index of performance, these assigned codes were converted to numeric values. Correct responses, assigned a value of 1, included "C", "P", "R", and "S"; incorrect responses, assigned a value of 0, included "M", "W", "N", and "F". The percentage of correct responses within the training and transfer segments of this task were computed as the summary indices of performance (for training, over 36 items; for transfer, over 57 items).

Over the three data collection years, two modifications were made in this task. First, the detail with which individual items were coded was decreased in scoring the Year 3 assessments (during the first two years, finer distinctions among incorrect responses were made). The second modification concerned administration procedures. As mentioned above, the training trials were stopped on a the student met a set performance criterion; however, for Years 1 and 2, students went through all training trials regardless of performance. For the Spanish version of the test, two items in the training materials were deleted from the analyses as their syllabic structure was found to lead to ambiguous segmentations.

Vocabulary. The student's knowledge of word meaning is assessed in three different formats in the SFST. The first two formats use pictures as vocabulary cues, while the third is entirely oral and entails no picture supports. The latter task was added after the Year 2 data collection period, and thus, only targets from Sites 3, 4, and 5 were tested with it.

The Fine Distinction section of this test assesses the student's ability to make fine distinctions between similar objects (nouns), actions (verbs), and relations (prepositions). In this task, closely related words are shown as picture triplets, and the student is asked to point to a specific item from each set (e.g., cake, pie, and doughnut are shown, and the student is asked to point to the pie). Vocabulary items included are those which represent concepts that are generally familiar to young school-age children (for documentation of the selection procedure, see Calfee, 1970). For each of the three form classes assessed, materials consist of four picture triplets, and each of the three items depicted is the proper response on a single occasion over the 12 stimulus presentations. Items are presented in a predetermined, random order. In this task, the pictures used in the English-language version are also used in the Spanish version, with appropriate translations of the verbal stimuli.

In the Common Label test, the student is first shown a set of pictures to name, and is then, for some items, asked to answer two qualitative questions which delve into other aspects of the concept depicted. The materials consist of 20 pictures, representing 10 nouns and 10 verbs. Approximately half the depicted words within each list are "common" words, and the rest are of "increasing difficulty," based on word lists acquired from multiple sources (for a full description of the selection procedure, see Calfee, 1970). For four of the words in each list, after the picture label has been given by the student, the questions "what is a ____ for?", and "what is a ____ made of?" are asked (for verbs, the two questions are "what is he ____ing?", and "why is he ____ing?").

For both the Fine Distinction and Common Label tasks, responses were coded as correct ("C"), incorrect ("W"), or no response ("N"). In deriving a summary index of performance, these assigned codes were converted to numeric values. Correct responses ("C") were assigned a value of 1, and incorrect responses ("W" and "N") were assigned a value of 0. A single summary measure was computed, representing the percent correct over the 72 items of the two tasks (Fine Distinction: 12 nouns, 12 verbs, and 12 prepositions; Common Label: 10 nouns, four with two probe questions each, and 10 verbs, four with two probe questions each).

Over the three years of administrations, no modifications were made to these two tasks. In the Spanish version, one item in the Fine Distinction list of nouns was deleted from all analyses as it was found to be a rare word for dialects common to the border sites.

The third test, Definitions, is entirely oral and entails no picture supports. The materials consist of three lists, each containing four words. Based on the rankings of Carroll, Davies, and Richman (1971), the words on List A have a frequency rank from 1 to 200, List B, from 201 to 300, and List C, from 301 to 450. For each word presented, the student is asked, "What does ____ mean?" If the student does not respond, then a second probe question is asked: "Can you think of another word that means the same as ____ ?" If there is still no response, or the student provides an incorrect definition, then the student is asked to try to select the correct definition from three alternatives spoken by the tester. The student is tested on each list in turn, from A through C, as long as at least half the words are correctly defined or the correct alternative is chosen.

The Spanish version employed most of the same words used in the English version for this task, but some words were not felt to be in the frequency range appropriate for a given list, and thus, were replaced with more appropriate words. Lacking the direction of any word frequency counts for Spanish usage in Mexico, these decisions were made based on the expert advice of the Spanish-language version collaboration team.

Each of the 12 items was coded as follows. (a) "C", a correct definition was given to either of the probe questions, (b) "M", a minimally adequate definition was given in response to the probe questions, (c) "P", a correct alternative was selected from the three alternatives given, (d) "W", no correct response was given to either the probe questions or the multiple-choice alternatives, (e) "N", no response was attained under any of the conditions, and (f) "F", assumed failure for words not tested because criterion on an earlier list was not attained. In deriving a summary index of performance, these assigned codes were converted to numeric values. Correct responses, assigned a value of 1, included "C", "M", and "P"; incorrect responses, assigned a value of 0, included "W", "N", and "F". The percent correct over the 12 items summarized performance on this task.

Story Comprehension. In this section of the SFST, the student is asked to listen to well-formed stories of increasing difficulty read by the tester, and to perform two comprehension tasks. In the first task, the student is asked to retell everything that can be remembered. In the second task, the tester asks probe questions about any major elements that are not completely recalled during the retelling. As with the Definition task described above, this task was added after the Year 2 data collection period, and thus, only targets from Sites 3, 4, and 5 were tested with it.

The materials for this task consist of stories which increase in both vocabulary difficulty and story complexity from Level A to Level C, with reading grade equivalents for the passages as follows: Level A, pre-primer to primer; Level B, grade 1; and Level C, grade 2. The Level A story contains four major story grammar elements; the Level B story, six elements; and the Level C story, eight elements. Associated with each of these elements is a probe question designed to elicit responses concerning the respective story element.

The Spanish version of each story followed the same story grammar sequence used in the English-language version. Each story was kept basically the same, but careful attention was given to assuring that the vocabulary of the stories was roughly equivalent to the grade level equivalents described above.

After a story was read by the tester and the student attempted to retell it, the tester asked the associated probe question for any element that was not completely given during free retell. The test was discontinued whenever the student failed to remember more than half of the elements of a given story under <u>either</u> free or cued recall procedures combined.

Each of the story elements under <u>free</u> recall was coded as follows: (a) "C", a complete and correct recall, (b) "B", a brief mention giving only partial information, (c) "N", no response for the element, and (d) "F", assumed failure, for story elements not attempted because the recall criterion on an earlier, less difficult story, was not attained. For <u>cued</u> recall, each element was coded as: (a) "P", a complete and correct recall, (b) "Q", a questionable recall, but with some relevance to the probed element, (c) "W", an incorrect response, (d) "N", no response, (e) "S", assumed success, for an element not probed because it was given completely and correctly during free retell, and (f) "F", assumed failure for story elements not probed because the recall criterion on an earlier story was not attained.

In deriving a summary index of performance, these assigned codes were converted to numeric values. Correct responses, assigned a value of 1, included "C" and "B" for free recall elements; "P", "Q" and "S" for cued recall elements. Incorrect responses, assigned a value of 0, included "N" and "F" for free recall elements; "W", "N", and "F" for cued recall. Since (a) the free recall score for a given element is partially embedded in the score for its associated cued recall, and (b) it was not felt that a fine level of detail was needed in capturing performance on this task, the percent correct over the 18 cued recall elements (four in Level A, six in Level B, and eight in Level C) was used as a summary index of performance in this task.

Reliability

The reliability assessments of the SFST summary measures are presented in Tables 2 and 3 (English and Spanish, respectively). These assessments were made for each of the data collection years, and therefore, the sample reflects the cohort structure discussed earlier (Volume 2: Design of the Study).

English administrations. For the English version, the students in the Year 1 and 2 samples represent an even mix of kindergarten and first-grade bilingual students, all from the border sites (the kindergarten students in the first cohort were tested again in the second year when they entered first grade). For Year 3, the sample consists entirely of kindergarten students from the two non-border sites; 20% are monolingual-English students and the remaining are bilingual.

As can be seen in Table 2, the reliability coefficients are all above .75, with the exception of the Year 3 assessment of Single-Letter Matching which has a coefficient of .69, indicating that the measures computed represent highly reliable summaries of performance in each of the tasks. More specific information will be given on the performance of the target sample in the next section, but a few comments about the performance of the total cohort are noteworthy -- remember, however, that there are site, language, and grade-level differences between cohorts, requiring caution in interpreting differences between testing years.

First, average performance across the cohorts is fairly constant, though there does appear to be some significant differences in variability. The first cohort, in general, outperforms the other two (this most likely is due to a self-selection artifact, since students during the first year were tested in the language version of their choice). Mean performance on Alphabet Knowledge is about 50%. Interestingly, the frequency distributions for this task (not shown) are bimodal -- a given student tends to either know all the letter names or none of them. Further, if the student is not successful on the production task, he tends also to be unsuccessful on the recognition task. Word Naming (only given in Year 3) shows that most students lack any sight-word recognition skill at entry to kindergarten (this distribution tends toward bimodality, but contains many more cases at the lower range than at the upper range). For the Letter Matching scales, performance is quite high, suggesting that most students come to school with sufficient skill, and do not require additional training in this area in preparation for reading instruction. For Auditory-Phonetic Segmentation, students in general master the training task. In fact, if one looks at the individual training trials, students tend to either master the task on the first trial, or require all of the trials in order to reach the performance criterion. In the Transfer section, performance is at about 50%, and an examination of the individual trials shows that performance is either high throughout the trials, or shows a steady decline as the materials become more unrelated to those employed in training. In looking at the last two transfer trials, the most difficult ones, the distributions are again distinctly bimodal.

Table 2

Stanford Foundation Skills Test - English:

Reliability Analysis of Total Scale Scores for Each Collection Year

Scale	Collection Year	N of Cases	N of Items*	Item Total Mean	Item Total SD	Mean Number of Actual Responses	α_k
Alphabet Knowledge	1	52	52	38.9	16.7	32.0	.98
	2	103	52	25.6	21.0	34.5	.99
	3	203	52	24.2	19.7	39.1	.98
Word Naming	3	202	12	0.9	1.9	6.1	.77
Single-Letter Matching	1	52	26	23.6	3.0	26.0	.80
	2	103	29	26.8	2.9	29.0	.80
	3	203	33	29.6	2.7	33.0	.59
Double-Letter Matching	1	52	35	30.4	4.9	35.0	.86
	2	103	36	29.8	5.9	36.0	.88
	3	203	36	27.2	6.6	36.0	.88
Auditory-Phonetic Segmentation: Training	1	52	33	30.4	3.1	33.0	.77
	2	102	36	31.8	6.8	36.0	.95
	3	202	36	34.0	5.8	11.7	.93
Auditory-Phonetic Segmentation: Transfer	1	52	56	43.1	10.4	56.0	.94
	2	102	57	34.3	14.2	57.0	.96
	3	202	57	35.3	16.6	55.0	.97
Vocabulary	1	52	55	45.5	5.2	55.0	.78
	2	103	72	50.8	14.5	72.0	.95
	3	200	72	58.4	11.1	72.0	.94
Definitions	3	200	12	6.0	4.5	9.5	.91
Comprehension	3	202	18	8.6	6.4	9.8	.90

*Items with no variance were deleted from the analysis. For each deleted item, its mean was 1.0 (i.e., all respondents answered correctly).

Note: For all scales allowing assumed success and failure, the reliability coefficient was adjusted for the number of such "responses" by reducing the residual degrees of freedom proportionately, and then recomputing the residual mean square and coefficient alpha on which each was based.

Table 3

Stanford Foundation Skills Test - Spanish:

Reliability Analysis of Total Scale Scores for Each Collection Year

Scale	Collection Year	N of Cases	N of Items*	Item Total Mean	Item Total SD	Mean Number of Actual Responses	αk
Alphabet Knowledge	1	248	60	10.6	14.4	57.9	.98
	2	102	60	10.1	12.2	73.7	.95
	3	212	60	10.9	14.3	47.3	.97
Word Naming	3	211	12	1.3	3.2	6.3	.93
Single-Letter Matching	1	247	35	30.9	4.5	35.0	.87
	2	103	24	21.9	2.7	24.0	.79
	3	219	33	30.2	2.7	33.0	.75
Double-Letter Matching	1	247	36	28.2	6.6	36.0	.89
	2	103	36	31.2	4.8	36.0	.85
	3	219	36	27.9	6.6	36.0	.78
Auditory-Phonetic Segmentation: Training	1	246	36	32.3	4.5	36.0	.86
	2	104	36	31.1	7.6	36.0	.95
	3	218	36	32.0	7.1	17.7	.91
Auditory-Phonetic Segmentation: Transfer	1	246	55	37.5	13.2	55.0	.96
	2	104	55	32.7	15.7	55.0	.97
	3	218	55	30.5	16.5	50.2	.97
Vocabulary	1	247	71	60.8	6.8	71.0	.85
	2	104	71	56.6	10.4	71.0	.93
	3	219	71	49.2	15.7	71.0	.96
Definitions	3	219	12	6.3	4.9	9.2	.94
Comprehension	3	218	18	7.7	7.2	8.6	.93

*Items with no variance were deleted from the analysis. For each deleted item, its mean was 1.0 (i.e., all respondents answered correctly).

Note: For all scales allowing assumed success and failure, the reliability coefficient was adjusted for the number of such "responses" by reducing the residual degrees of freedom proportionately, and then recomputing the residual mean square and coefficient alpha on which each was based.

In the Vocabulary task, average performance is fairly high at about 70%. In the remaining language tasks (Definitions and Comprehension, given only in Year 3), students are successful on about half of the material sets; again, these distributions are bimodal.

Spanish administrations. Turning to the Spanish reliability analyses summarized in Table 3, one can see that a much larger sample was tested in Year 1 as compared to the English sample. This was done in order to allow a sufficient assessment of the newly-developed Spanish version of the instrument; all students tested were from Site 0, with an even mix of kindergarten and first-grade bilingual students (the target sample comprised about 20% of this group). In Year 2, the sample consisted of the same students assessed with the English version (i.e., an even mix of bilingual kindergarten and first-grade students, all from the border sites). The Year 3 sample contained the same bilingual students that were assessed with the English version in this year (i.e., all kindergarten students from the non-border sites); however, it contained no monolingual-English students, but did include the monolingual-Spanish first-grade targets from Site 4 (representing about 25% of this sample).

Keeping in mind the cautions given above about comparisons between cohorts, the following statements can be made. First, each of the reliability coefficients is quite high, and average performance is fairly constant across cohorts, though again, there appear to be some significant differences in variability. Alphabet Knowledge is, on average, negligible. This is not surprising when one considers that Spanish-speaking children generally are introduced to the alphabet through the sounds of the letters rather than the letter names. In fact, letter names are infrequent in children's literature in Spanish except for the vowels, in which case their sounds and letter names coincide. In English, children's literature is replete with ABC songs and rhymes, thus providing many opportunities for English-speaking preschool children to learn the names of the letters. Accordingly, the frequency distributions for both of the Alphabet Knowledge tasks are unimodal (unlike those in English), with a preponderance of cases at the lower range. In Word Naming, most students are unable to recognize any of the items presented; the distribution is quite similar to the one obtained from the English testing -- most cases at the bottom range, but some cases revealing successful recognition of all items. In Letter Matching, there is a high level of performance, as was found in English, which, given the non-verbal nature of the task, would be expected, assuming no language difficulties occur in understanding the tester's instructions. Also, in Auditory-Phonetic Segmentation, both the average performance, as well as the shape of the distributions, are similar to those found in English. Finally, performance in Definitions and Comprehension is at the level found in English (about 50% in the aggregate), and the distributions show the same bimodalities.

Descriptive Statistics

In this section, the performance of the bilingual sample is described, first for the English version, then for the Spanish version.

English administrations. The English data for this sample are presented in Tables 4 through 10. As each of these are organized in a similar fashion, the layout of the first table will be given in some detail.

In Table 4, the left margin is defined by the nine tasks, and for each, the mean (M), standard deviation (S), and number of cases (N) are given for every category appearing along the top of the table. The task names are mnemonic, and stand for the following scale names:

ALPHPR: Alphabet Knowledge (Production - Recognition)
WRDNAM: Word Naming
SNGLTM: Single-Letter Matching
DBLLTM: Double-Letter Matching
PSTRNG: Auditory-Phonetic Segmentation - Training
PSTRNF: Auditory-Phonetic Segmentation - Transfer
VCFDCL: Vocabulary (Fine Distinctions - Common Labels)
DEFNTN: Definitions
CMPPRB: Comprehension (Probes)

The two letters appended to each of the task names give the language of the test, as English or Spanish (E or S, respectively), and the grade level of administration, as kindergarten or first (K or F, respectively).

Along the top of the table, the first column provides data based on the entire bilingual sample (Overall), and then successively for students in the language entry categories of low English (Low Eng), high English (High Eng), low Spanish (Low Span), and high Spanish (High Span). These are then followed by a further refinement of language category based on combined English and Spanish entry skill: low English and low Spanish (Lo Lo), low English and high Spanish (Lo Hi), high English and low Spanish (Hi Lo), and high English and high Spanish (Hi Hi). These language categories have been described elsewhere in detail (see Volume 4: Oral Language Growth), but as a review, they reflect a division of the targets based on teacher ratings of English and Spanish skill at entry to kindergarten. The rating scale consisted of 5 points, from low to high, and in an effort to achieve an approximate even distribution of students within the two languages, a value of 3.0 or above was used for the high English category, and 4.0 or above for the high Spanish category.

The other tables follow a similar pattern: Table 5 provides the overall data for the English first-grade administration, and Tables 6 through 10 provide individual site data (Sites 0, 1, 2, 3, and 5, respectively) with kindergarten data in the top panel, and first-grade data (if obtained) in the bottom panel. For these last five tables, the overall site data are provided first (Overall), followed by the data for the four language entry categories (Lo Lo, Lo Hi, Hi Lo, and Hi Hi). The number appended to each of these labels is simply the site identification number.

Table 4

Stanford Foundation Skills Test - English:
Descriptive Statistics for Each Scale for Kindergarten Bilingual Sample
Overall and by Language Category

Scale	Statistic	Overall	Low Eng	High Eng	Low Span	High Span	Lo Lo	Lo Hi	Hi Lo	Hi Hi
ALPHPR-EK	M	40.4	23.2	53.5	36.5	44.3	26.1	19.5	46.0	60.0
ALPHPR-EK	S	36.5	27.6	37.2	34.8	38.1	28.2	26.8	37.7	35.8
ALPHPR-EK	N	162	70	92	82	80	39	31	43	49
WRDNAM-EK	M	6.8	1.7	10.3	6.6	7.1	1.2	2.8	11.2	9.1
WRDNAM-EK	S	14.7	5.3	17.8	15.3	13.8	4.1	7.5	19.5	15.6
WRDNAM-EK	N	122	50	72	75	47	35	15	40	32
SNGLTM-EK	M	90.0	87.7	91.7	88.8	91.2	87.4	88.0	90.1	93.2
SNGLTM-EK	S	8.5	8.8	7.9	8.8	8.1	8.5	9.4	8.9	6.5
SNGLTM-EK	N	162	70	92	82	80	39	31	43	49
DBLLTM-EK	M	76.7	73.3	79.3	75.8	77.6	75.6	70.3	76.0	82.1
DBLLTM-EK	S	18.6	18.7	18.2	18.0	19.3	17.3	20.2	18.7	17.3
DBLLTM-EK	N	162	70	92	82	80	39	31	43	49
PSTRNG-EK	M	93.2	88.2	96.9	96.0	90.2	92.4	82.9	99.3	94.9
PSTRNG-EK	S	16.2	20.9	9.9	11.4	19.6	15.7	25.3	1.7	13.2
PSTRNG-EK	N	162	70	92	82	80	39	31	43	49
PSTRNF-EK	M	60.0	49.6	68.0	61.8	58.2	54.5	43.5	68.5	67.5
PSTRNF-EK	S	28.1	27.3	26.1	27.8	28.4	28.0	25.4	26.2	26.3
PSTRNF-EK	N	162	70	92	82	80	39	31	43	49
VCFDCL-EK	M	75.2	63.6	84.0	78.4	71.8	72.5	52.1	83.8	84.1
VCFDCL-EK	S	19.3	23.0	9.0	14.5	23.0	18.0	23.9	7.0	10.[illegible]
VCFDCL-EK	N	160	69	91	82	78	39	30	43	48
DEFNTN-EK	M	43.5	27.2	54.9	39.2	50.6	24.1	35.1	52.9	57.5
DEFNTN-EK	S	35.4	32.7	32.9	34.2	36.7	29.7	39.5	32.5	33.7
DEFNTN-EK	N	121	50	71	76	45	36	14	40	31
CMPPRB-EK	M	42.5	24.7	55.2	40.6	45.6	27.0	19.3	52.9	58.0
CMPPRB-EK	S	34.0	29.1	31.6	31.1	38.3	28.8	30.3	28.2	35.6
CMPPRB-EK	N	123	51	72	76	47	36	15	40	32

Table 5

Stanford Foundation Skills Test - English:
Descriptive Statistics for Each Scale for First-grade Bilingual Sample
Overall and by Language Category

Scale	Statistic	Overall	Low Eng	High Eng	Low Span	High Span	Lo Lo	Lo Hi	Hi Lo	Hi Hi
ALPHPR-EF	M	64.3	47.6	80.9	62.0	65.2	47.3	47.8	98.6	77 6
ALPHPR-EF	S	34.9	32.3	29.4	40.4	33.1	38.9	28.5	1.8	31.1
ALPHPR-EF	N	50	25	25	14	36	10	15	4	21
WRDNAM-EF	M									
WRDNAM-EF	S									
WRDNAM-EF	N									
SNGLTM-EF	M	95.9	94.7	97.0	92.5	97.2	91.4	96.9	95.0	97.4
SNGLTM-EF	S	6.8	7.3	6.1	8.3	5.7	9.3	4.8	4.9	6.3
SNGLTM-EF	N	50	25	25	14	36	10	15	4	21
DBLLTM-EF	M	88.1	83.1	93.0	87.5	88.3	83.9	82.6	96.5	92.3
DBLLTM-EF	S	13.1	14.4	9.5	12.7	13.4	13.1	15.6	5.3	10.1
DBLLTM-EF	N	50	25	25	14	36	10	15	4	21
PSTRNG-EF	M	94.8	93.3	96.2	93.1	95.4	90.3	95.4	100.0	95.5
PSTRNG-EF	S	10.1	13.1	5.6	15.8	6.9	18.2	8.3	0.0	5.9
PSTRNG-EF	N	50	25	25	14	36	10	15	4	21
PSTRNF-EF	M	69.4	61.8	77.1	69.1	69.6	63.7	60.5	82.5	76.1
PSTRNF-EF	S	20.1	19.7	17.6	23.7	18.8	25.3	15.9	13.8	18.3
PSTRNF-EF	N	50	25	25	14	36	10	15	4	21
VCFDCL-EF	M	76.9	68.7	85.0	73.9	78.0	68.6	68.8	87.2	84.6
VCFDCL-EF	S	13.3	11.8	9.3	16.1	12.2	16.1	8.4	4.7	10.0
VCFDCL-EF	N	50	25	25	14	36	10	15	4	21
DEFNTN-EF	M									
DEFNTN-EF	S									
DEFNTN-EF	N									
CMPPRB-EF	M									
CMPPRB-EF	S									
CMPPRB-EF	N									

287

Table 6

Stanford Foundation Skills Test - English:
Descriptive Statistics for Each Scale for Site 0 Bilingual Sample
Overall and by Language Category

Scale	Statistic	Overall	Lo Lo-0	Lo Hi-0	Hi Lo-0	Hi Hi-0
ALPHPR-EX	M	30.8	0.0	2.2		37.9
ALPHPR-EX	S	43.7	0.0	2.3		46.4
ALPHPR-EX	N	16	2	6		8
WRDNAM-EX	M					
WRDNAM-EX	S					
WRDNAM-EX	N					
SNGLTR-EX	M	94.1	95.7	88.1		98.2
SNGLTR-EX	S	7.0	6.1	8.0		2.1
SNGLTR-EX	N	16	2	6		8
DBLLTR-EX	M	82.1	66.7	73.6		92.4
DBLLTR-EX	S	18.5	39.2	16.7		8.1
DBLLTR-EX	N	16	2	6		8
PSTRNB-EX	M	85.5	70.9	69.9		96.9
PSTRNB-EX	S .	17.9	21.6	16.5		4.1
PSTRNB-EX	N	16	2	6		8
PSTRNF-EX	M	60.0	46.5	41.2		77.4
PSTRNF-EX	S	22.8	26.0	12.8		14.0
PSTRNF-EX	N	16	2	6		8
VCFBCL-EX	M	58.1	31.3	35.2		82.0
VCFBCL-EX	S	26.3	14.8	8.8		9.3
VCFBCL-EX	N	16	2	6		8
DEFNTN-EX	M					
DEFNTN-EX	S					
DEFNTN-EX	N					
CNPPRD-EX	M					
CNPPRD-EX	S					
CNPPRD-EX	N					
ALPHPR-EF	M	57.8	39.4	50.0		68.4
ALPHPR-EF	S	33.3	31.3	23.2		34.7
ALPHPR-EF	N	22	4	4		14
WRDNAM-EF	M					
WRDNAM-EF	S					
WRDNAM-EF	N					
SNGLTR-EF	M	98.2	95.7	97.9		99.0
SNGLTR-EF	S	3.0	5.5	2.7		1.8
SNGLTR-EF	N	22	4	4		14
DBLLTR-EF	M	95.8	91.7	95.1		97.2
DBLLTR-EF	S	4.3	3.1	4.7		3.3
DBLLTR-EF	N	22	4	4		14
PSTRNB-EF	M	89.9	80.6	87.5		93.2
PSTRNB-EF	S	13.4	27.5	13.1		6.0
PSTRNB-EF	N	22	4	4		14
PSTRNF-EF	M	74.6	66.2	64.9		79.7
PSTRNF-EF	S	19.8	27.1	16.7		10.0
PSTRNF-EF	N	22	4	4		14
VCFBCL-EF	M	77.0	66.0	67.0		83.1
VCFBCL-EF	S	14.9	22.1	5.2		11.5
VCFBCL-EF	N	[illegible]	4	4		14
DEFNTN-EF	M					
DEFNTN-EF	S					
DEFNTN-EF	N					
CNPPRD-EF	M					
CNPPRD-EF	S					
CNPPRD-EF	N					

288

Table 7

Stanford Foundation Skills Test - English:
Descriptive Statistics for Each Scale for Site Bilingual Sample
Overall and by Language Category

Scale	Statistic	Overall	Lo Lo-1	Lo Hi-1	Hi Lo-1	Hi Hi-1
ALPHPR-EX	M	56.9		38.0	82.7	81.7
ALPHPR-EX	S	37.5		41.1	0.0	6.8
ALPHPR-EX	N	7		4	1	2
WRDRCN-EX	M					
WRDRCN-EX	S					
WRDRCN-EX	N					
SNGLTR-EX	M	92		89	97	97
SNGLTR-EX	S	7.5		8.4	6.0	0.0
SNGLTR-EX	N	7		4	1	2
DBLLTR-EX	M	84.5		75.7	94.4	97.2
DBLLTR-EX	S	20		23.9	0.0	4.0
DBLLTR-EX	N	7		4	1	2
PSTRNB-EX	M	77.8		63.3	91.7	95.9
PSTRNB-EX	S	30.9		37.6	0.0	5.9
PSTRNB-EX	N	7		4	1	2
PSTRNF-EX	M	44.6		39.5	36.8	58.8
PSTRNF-EX	S	23.2		26.6	0.0	23.5
PSTRNF-EX	N	7		4	1	2
VCFBCL-EX	M	67.5		56.6	70.8	87.5
VCFBCL-EX	S	23.3		25.6	0.0	[illegible]
VCFBCL-EX	N	7		4	1	2
DEFNTN-EX	M					
DEFNTN-EX	S					
DEFNTN-EX	N					
CMPPRD-EX	M					
CMPPRD-EX	S					
CMPPRD-EX	N					
ALPHPR-EF	M	69.0	56.3	51.0	97.2	100.0
ALPHPR-EF	S	39.1	48.4	36.7	1.3	0.0
ALPHPR-EF	N	9	4	2	2	1
WRDRCN-EF	M					
WRDRCN-EF	S					
WRDRCN-EF	N					
SNGLTR-EF	M	93.0	80.6	96.6	92.9	100.0
SNGLTR-EF	S	8.2	10.2	2.1	6.0	0.0
SNGLTR-EF	N	9	4	2	2	1
DBLLTR-EF	M	91.7	83.3	97.2	98.6	100.0
DBLLTR-EF	S	8.6	4.5	4.0	2.0	0.0
DBLLTR-EF	N	9	4	2	2	1
PSTRNB-EF	M	96.6	95.1	94.5	100.0	100.0
PSTRNB-EF	S	4.8	4.7	7.8	0.0	0.0
PSTRNB-EF	N	9	4	2	2	1
PSTRNF-EF	M	72.3	65.8	61.4	86.0	93.0
PSTRNF-EF	S	20.8	25.4	4.9	17.3	0.0
PSTRNF-EF	N	9	4	2	[illegible]	1
VCFBCL-EF	M	76.3	73.6	66.7	86.2	86.1
VCFBCL-EF	S	11.8	10.2	15.7	7.0	0.0
VCFBCL-EF	N	9	4	2	2	1
DEFNTN-EF	M					
DEFNTN-EF	S					
DEFNTN-EF	N					
CMPPRD-EF	M					
CMPPRD-EF	S					
CMPPRD-EF	N					

25[illegible]

Table 8

Stanford Foundation Skills Test - English:
Descriptive Statistics for Each Scale for Site 2 Bilingual Sample
Overall and by Language Category

Scale	Statistic	Overall	Lo Lo-2	Lo Hi-2	Hi Lo-2	Hi Hi-2
ALPHPR-EK	M	32.1	0.0	16.6	12.5	62.4
ALPHPR-EK	S	37.8	0.0	29.7	1.4	35.4
ALPHPR-EK	N	16	1	6	2	7
WRDRAN-EK	M					
WRDRAN-EK	S					
WRDRAN-EK	N					
SNGLTR-EK	M	87	86	81	74	95
SNGLTR-EK	S	12.4	0.0	11.8	4.0	10.5
SNGLTR-EK	N	16	1	6	2	7
DBLLTR-EK	M	74.1	100.0	62.0	66.7	82.9
DBLLTR-EK	S	16.0	0.0	11.3	16	11.4
DBLLTR-EK	N	16	1	6	2	7
PSTRNB-EK	M	80.4	61.1	69.0	100.0	87.3
PSTRNB-EK	S	23.5	0.0	33.9	0.0	13.0
PSTRNB-EK	N	16	1	6	2	7
PSTRNF-EK	M	37.8	22.8	23.1	30.8	54.6
PSTRNF-EK	S	23.0	0.0	20.9	13.6	18.1
PSTRNF-EK	N	16	1	6	2	7
VCFBCL-EK	M	67.4	79.2	43.1	81.3	82.7
VCFBCL-EK	S	23.1	0.0	17.9	16.8	7.8
VCFBCL-EK	N	16	1	6	2	7
DEFNTN-EK	M					
DEFNTN-EK	S					
DEFNTN-EK	N					
CMPPRD-EK	M					
CMPPRD-EK	S					
CMPPRD-EK	N					
ALPHPR-EF	M	67.2	45.2	46.2	100.0	95.2
ALPHPR-EF	S	36.0	35.7	32.3	0.0	5.4
ALPHPR-EF	N	19	2	9	2	6
WRDRAN-EF	M					
WRDRAN-EF	S					
WRDRAN-EF	N					
SNGLTR-EF	M	94.6	88.6	96.2	97.2	93.3
SNGLTR-EF	S	8.5	16.2	5.9	4.0	11.1
SNGLTR-EF	N	19	2	9	2	6
DBLLTR-EF	M	77.3	69.5	73.8	94.5	79.6
DBLLTR-EF	S	14.5	27.5	14.0	7.8	10.3
DBLLTR-EF	N	19	2	9	2	6
PSTRNB-EF	M	99.6	100.0	99.1	100.0	100.0
PSTRNB-EF	S	1.0	0.0	1.4	0.0	0.0
PSTRNB-EF	N	19	2	9	2	6
PSTRNF-EF	M	62.1	54.4	58.3	77.0	64.9
PSTRNF-EF	S	18.7	37.2	17.9	14.9	16.0
PSTRNF-EF	N	19	2	9	2	6
VCFBCL-EF	M	77.0	63.9	70.1	88.2	88.0
VCFBCL-EF	S	12.8	19.7	8.9	1.0	5.8
VCFBCL-EF	N	19	2	9	2	6
DEFNTN-EF	M					
DEFNTN-EF	S					
DEFNTN-EF	N					
CMPPRD-EF	M					
CMPPRD-EF	S					
CMPPRD-EF	N					

Table 9

Stanford Foundation Skills Test - English:
Descriptive Statistics for Each Scale for Site 3 Bilingual Sample
Overall and by Language Category

Scale	Statistic	Overall	Lo L .	Lo Hi-3	Hi Lo-3	Hi Hi-3
ALPHPR-EK	M	40.0	29.6		45.7	46.2
ALPHPR-EK	S	35.4	30.6		38.0	34.3
ALPHPR-EK	N	67	24		35	8
WRDNAM-EK	M	8.8	1.7		11.9	16.7
WRDNAM-EK	S	17.5	4.9		20.4	22.3
WRDNAM-EK	N	67	24		35	8
SNGLTM-EK	M	88.7	85.2		90.5	91.1
SNGLTM-EK	S	8.6	8.3		8.9	4.9
SNGLTM-EK	N	67	24		35	8
DBLLTM-EK	M	77.8	74.9		79.0	81.6
DBLLTM-EK	S	16.4	16.2		17.1	14.0
DBLLTM-EK	N	67	24		35	8
PSTRNG-EK	M	96.9	92.5		99.4	99.3
PSTRNG-EK	S	10.5	16.8		1.3	2.0
PSTRNG-EK	N	67	24		35	8
PSTRNF-EK	M	64.2	52.4		71.6	67.1
PSTRNF-EK	S	26.5	25.4		25.4	21.1
PSTRNF-EK	N	67	24		35	8
VCFDCL-EK	M	82.8	78.5		84.7	87.5
VCFDCL-EK	S	8.0	9.5		5.6	6.6
VCFDCL-EK	N	67	24		35	8
DEFNTN-EK	M	39.0	24.6		49.5	36.5
DEFNTN-EK	S	34.4	29.7		33.2	40.3
DEFNTN-EK	N	67	24		35	8
CMPPRB-EK	M	45.8	32.2		53.3	53.5
CMPPRB-EK	S	29.9	30.5		25.7	34.0
CMPPRB-EK	N	67	24		35	8

291

Table 10

Stanford Foundation Skills Test - English:
Descriptive Statistics for Each Scale for Site 5 Bilingual Sample
Overall and by Language Category

Scale	Statistic	Overall	Lo Lo-5	Lo Hi-5	Hi Lo-5	Hi Hi-5
ALPHPR-EK	M	43.1	25.6	22.6	54.6	62.3
ALPHPR-EK	S	35.5	23.8	25.0	40.9	34.9
ALPHPR-EK	N	56	12	15	5	24
WRDNAM-EK	M	4.2	0.0		6.7	6.6
WRDNAM-EK	S	9.7	0.0		10.9	12.3
WRDNAM-EK	N	55	11		5	24
SNGLTM-EK	M	91.0	90.5	90.5	92.0	91.[illegible]
SNGLTM-EK	S	7.1	8.3	8.7	4.7	6.0
SNGLTM-EK	N	56	12	15	5	24
DBLLTM-EK	M	73.5	76.4	70.9	55.0	77.4
DBLLTM-EK	S	21.2	16.7	23.7	19.6	20.8
DBLLTM-EK	N	56	12	15	5	24
PSTRNG-EK	M	97.0	98.6	98.3	99.4	94.9
PSTRNG-EK	S	11.3	2.2	3.9	1.3	16.9
PSTRNG-EK	N	56	12	15	5	24
PSTRNF-EK	M	63.4	62.7	53.6	68.1	68.9
PSTRNF-EK	S	30.2	31.9	26.9	25.1	32.4
PSTRNF-EK	N	56	12	15	5	24
VCFDCL-EK	M	74.2	66.9	62.0	81.7	83.8
VCFDCL-EK	S	21.1	22 1	26.2	10.7	13.0
VCFDCL-EK	N	54	12	14	5	23
DEFNTN-EK	M	48.9	22.9	35.1	76.7	64.8
DEFNTN-EK	S	36.3	30.8	39.5	12.4	28.5
DEFNTN-EK	N	54	12	14	5	23
CMPPRB-EK	M	38.7	16.7	19.3	50.0	59.5
CMPPRB-EK	S	38.2	22.6	30.3	45.8	36.7
CMPPRB-EK	N	56	12	15	5	24

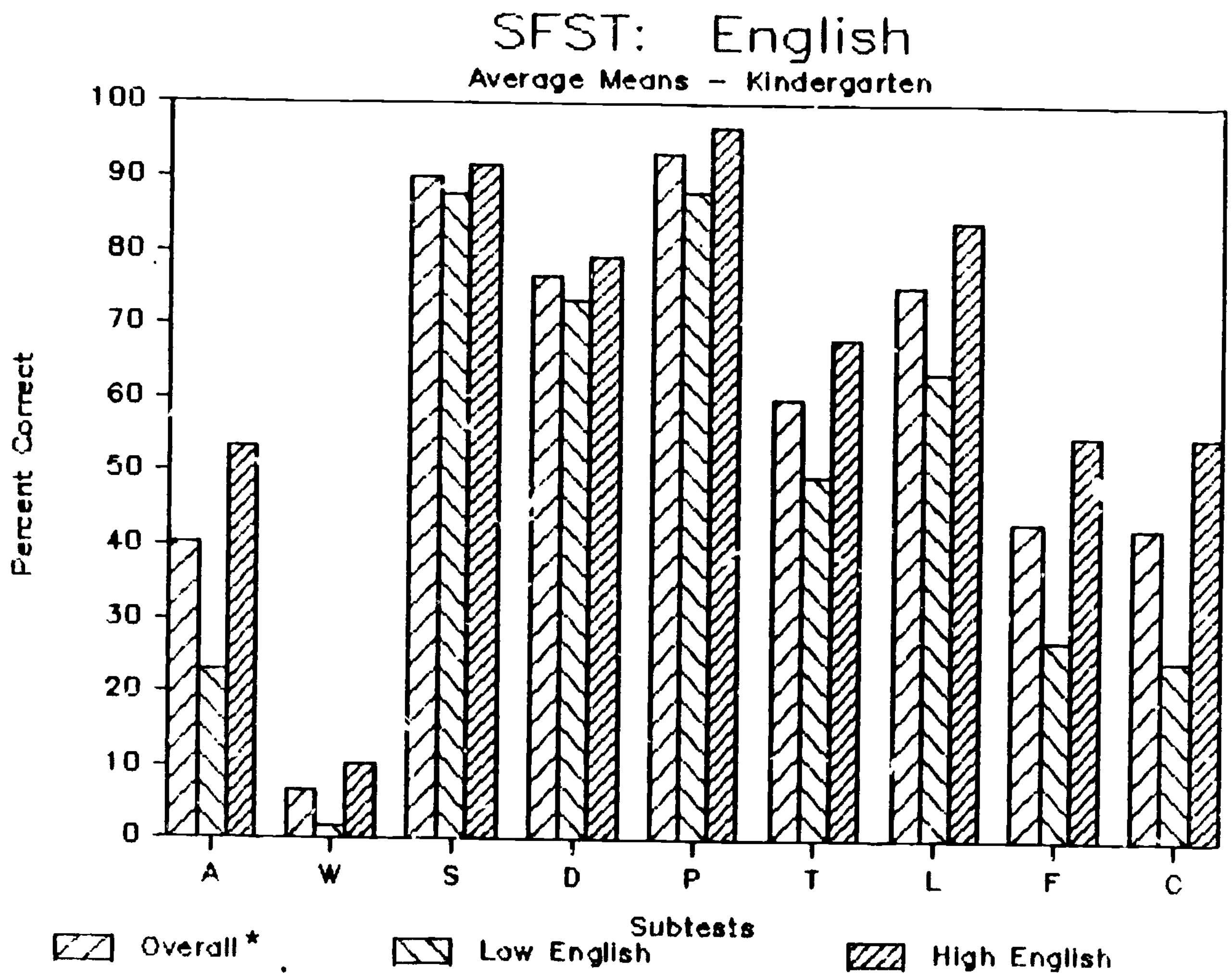
When considering these data, it is important to keep in mind the distribution of administrations by site and grade level presented in Table 1. All sites are represented in the kindergarten data (though only about half of the border site target students were tested then), while at first grade, only the border sites (0, 1, and 2) are represented. Further, for the tasks added in the Year 3 administrations (Word Naming, Definitions, and Comprehension), only the non-border sites (3 and 5) are represented.

In Figure 1, the average performance for the nine SFST tasks is presented over the entire sample of bilingual students tested in kindergarten, as well as performance broken out on the basis of the English entry categories (the actual values come from Table 4). As can be seen, performance generally matches that described above for the individual cohorts of students. Average performance on Alphabet Knowledge is at about 40%. Little skill is shown in Word Naming, and Single- and Double-Letter Matching performance is high (though the latter is found to be more difficult than the former). Most students master the Auditory-Phonetic Segmentation Training task, but some have difficulty in transferring the skill to novel materials. Performance on the Vocabulary task is at about 70%, with Definition and Comprehension skills lower at about 40%. The high English students tend to outperform the low English students, and without assessing these differences through the usual statistical tests, it seems that the most substantial differences are reflected mainly in the language dependent tasks of Alphabet Knowledge, Vocabulary, Definitions, and Comprehension, with little differences in the non-linguistic tasks of Letter Matching. Interestingly, Auditory-Phonetic Segmentation seems to be acquired equally well by both groups, but the high English group shows a small advantage in transferring this skill to new material, consisting of mostly real English words in the early trials, and synthetic words in the final trials.

Figure 2 presents the same data for the combined English-Spanish categories. One can detect little difference within the high English groups based on their differing Spanish skills. Within the low English groups, the trend appears to be that the low Spanish group outperforms the high Spanish group in the linguistic tasks, but with no differences in the non-linguistic tasks.

Figure 3 displays the kindergarten English data for each site (the data are taken from Tables 6 through 10). There appears to be some substantial site differences in the Auditory-Phonetic Segmentation tasks and the Vocabulary task, with the non-border sites outperforming the border sites -- no systematic trends appear in the other scales. However, in examining the data from the individual sites, the differences between low and high English categories for each of the scales in general follow the descriptions given above for the overall data set.

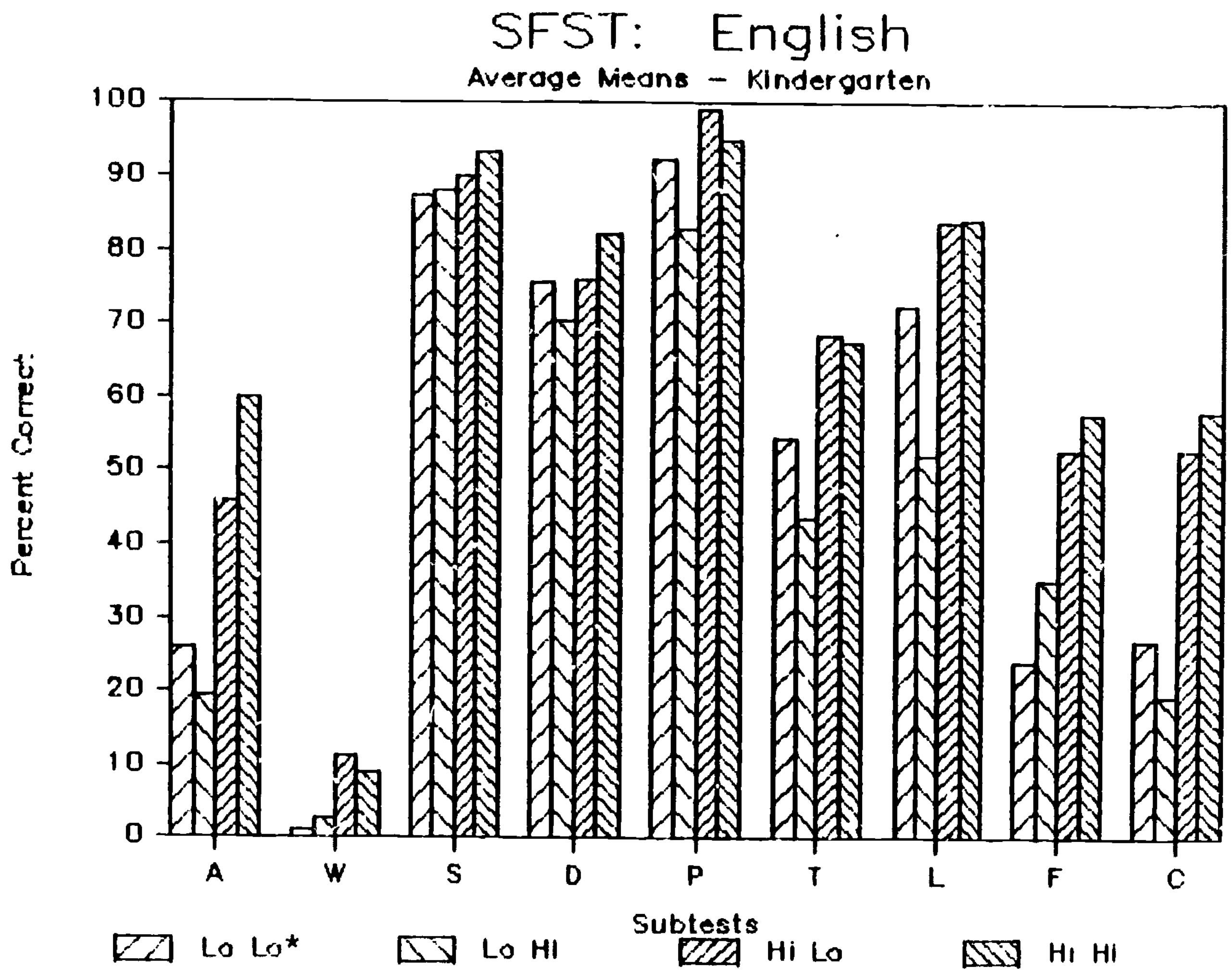
Figures 4 and 5 present the first-grade data obtained for those border site students who entered the study as first-graders. Again, performance follows that described for the individual cohorts, and, in general, is higher than that obtained from the kindergarten sample.



A = Alphabet Knowledge
W = Word Naming
S = Single-Letter Matching
D = Double-Letter Matching
P = Auditory-Phonetic Segmentation: Training
T = Auditory-Phonetic Segmentation: Transfer
L = Fine Distinctions Common Labels
F = Definitions
C = Comprehension

*N=122 for the W, F, and C scales; otherwise N=162.

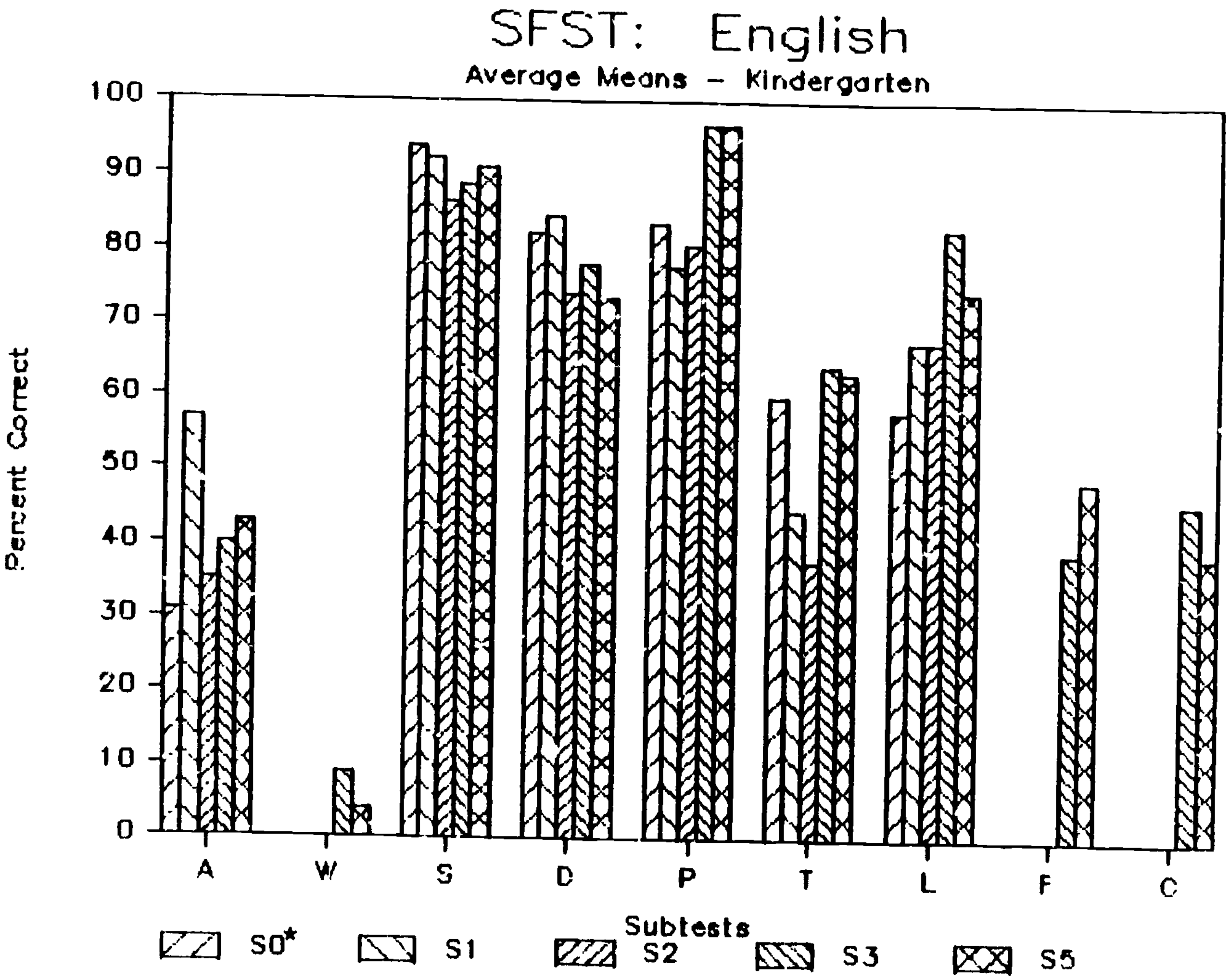
Figure 1. SFST-E scale means for the kindergarten entry bilingual sample overa.. and by English entry category.



A = Alphabet Knowledge
W = Word Naming
S = Single-Letter Matching
D = Double-Letter Matching
P = Auditory-Phonetic Segmentation: Training
T = Auditory-Phonetic Segmentation: Transfer
L = Fine Distinctions Common Labels
F = Definitions
C = Comprehension

*Low English, Low Spanish entry.

Figure 2. SFST-E scale means for the kindergarten entry bilingual sample by English-Spanish entry category.



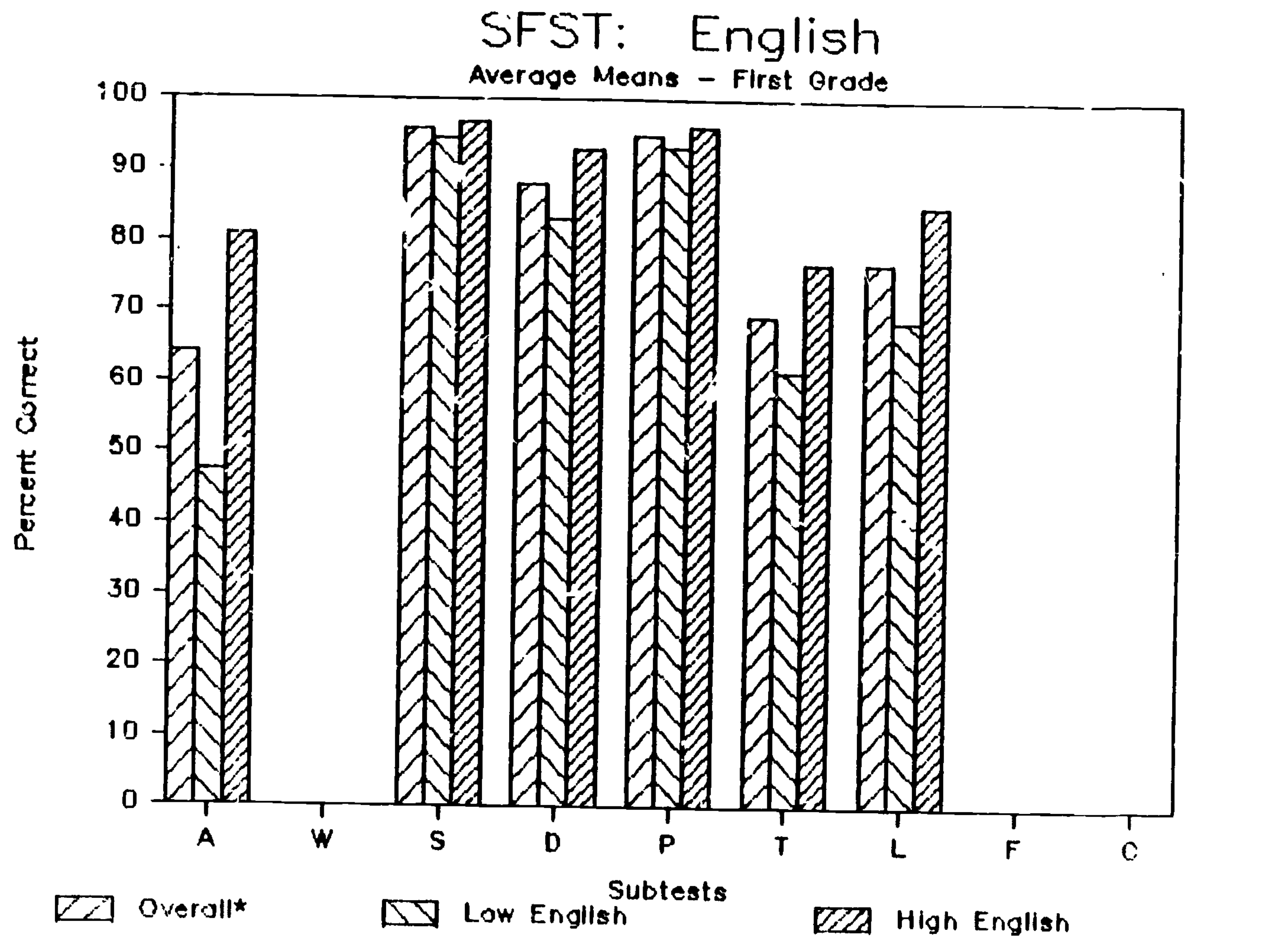
A = Alphabet Knowledge
W = Word Naming
S = Single-Letter Matching

D = Double-Letter Matching
P = Auditory-Phonetic Segmentation: Training
T = Auditory-Phonetic Segmentation: Transfer

L = Fine Distinctions Common Labels
F = Definitions
C = Comprehension

*Site 0.

Figure 3. SFST-E scale means for the kindergarten entry bilingual sample by site.



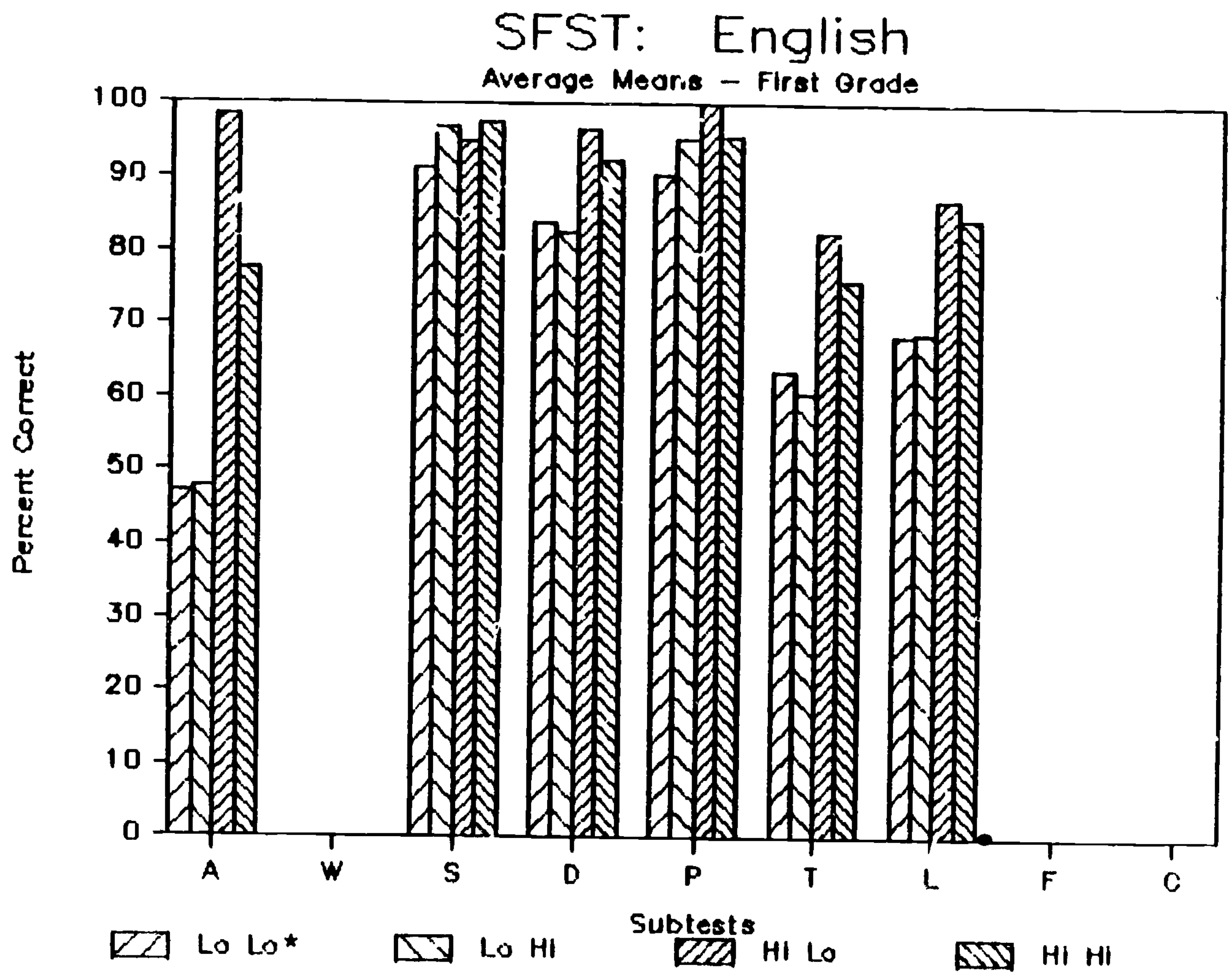
A = Alphabet Knowledge
W = Word Naming
S = Single-Letter Matching

D = Double-Letter Matching
P = Auditory-Phonetic Segmentation: Training
T = Auditory-Phonetic Segmentation: Transfer

L = Fine Distinctions Common Labels
F = Definitions
C = Comprehension

*N=50.

Figure 4. SFST-E scale means for the first-grade entry bilingual sample overall and by English entry category.



A = Alphabet Knowledge
W = Word Naming
S = Single-Letter Matching

D = Double-Letter Matching
P = Auditory-Phonetic Segmentation: Training
T = Auditory-Phonetic Segmentation: Transfer

L = Fine Distinctions Common Labels
F = Definitions
C = Comprehension

*Low English, Low Spanish entry.

Figure 5. SFST-E scale means for the first-grade entry bilingual sample by English-Spanish entry category.

Further, the same relationship between the low and high English groups holds: higher performance for the latter in the linguistic tasks, no apparent differences in the non-linguistic tasks.

Spanish administrations. The Spanish data are presented in Tables 11 through 17. The formats of these tables are identical to those used for the English data (see the description above). Table 11 provides the kindergarten data for the entire sample, along with the breakdowns by language category. This is followed by Table 12 which gives the first-grade data. Tables 13 through 17 present the data for each site (0, 1, 2, 3, and 5, respectively), giving the overall site avarages followed by breakdowns by the combined English-Spanish categories.

As cautioned above, when considering these data, it is important to keep in mind the distribution of administrations by site and grade level presented in Table 1. Again, all sites are represented in the kindergarten data (though only about half of the border site students were tested then), while at first grade, only the border sites (0, 1, and 2) are represented. Further, for the tasks added in the Year 3 administrations (Word Naming, Definitions, and Comprehension), only the non-border sites (3 and 5) are represented.

In Figure 6 (based on Table 11), the Spanish kindergarten entry data are displayed for the overall sample, and for the component low and high Spanish entry groups. These data, as found in the English set, follow the same general patterns described above for the individual cohorts. First, knowledge of the alphabet is negligible, as is sight-word recognition. Letter Matching skill is high, though the Double-Letter task presents more difficulty than the Single-Letter task. The Auditory-Phonetic Segmentation Training trials are successfully acquired by most students, but they seem to have considerable difficulty with the Transfer materials. The Vocabulary task performance is high with an average of about 80%, with the performance on the Definitions and Comprehension tasks somewhat lower. Again, without resort to the standard statistical tests, the high Spanish group seems to outperform the low Spanish group on the linguistic tasks, with no apparent difference in the non-linguistic tasks. Although the English and Spanish samples are not strictly comparable, in general, linguistic task performance appears to be greater in Spanish than in English.

Figure 7 displays the Spanish kindergarten data by the combined English-Spanish entry skill categories. As before, performance for both of the high Spanish groups seems to exceed that of the two low Spanish groups on the linguistic tasks. Further, for the low Spanish groups, there do not appear to be any substantial differences except for the last two tasks (Definitions and Comprehension), where the low English group outperforms the low Spanish group -- remember that these data come exclusively from the non-border Sites 3 and 5.

Figure 8 charts the Spanish kindergarten data for each site (the data come from Tables 13 through 17). There appear to be no systematic trends with the following exceptions: (a) Auditory-Phonetic Segmentation performance appears to be greater in Sites 0 and 5 for

Table 11

St... Foundation Skills Test - Spanish:
Descriptive Statistics for Each Scale for Kindergarten Bilingual Sample
Overall and by Language Category

Scale	Statistic	Overall	Low Eng	High Eng	Low Span	High Span	Lo Lo	Lo Hi	Hi Lo	Hi Hi
ALPHPR-SK	M	9.3	6.7	11.4	7.9	10.6	6.7	6.8	9.1	13.1
ALPHPR-SK	S	13.1	11.0	14.3	12.8	13.4	13.2	7.5	12.3	15.6
ALPHPR-SK	N	179	80	99	87	92	44	36	43	56
WRDNAM-SK	M	1.9	1.4	2.2	1.5	2.5	1.3	1.7	1.7	2.9
WRDNAM-SK	S	5.2	5.4	5.0	5.2	5.2	5.7	4.7	4.7	5.4
WRDNAM-SK	N	125	53	72	78	47	38	15	40	32
SNGLTM-SK	M	90.9	88.9	92.5	91.4	90.4	90.1	87.5	92.7	92.3
SNGLTM-SK	S	7.7	7.5	7.5	7.0	8.3	6.1	8.8	7.6	7.5
SNGLTM-SK	N	180	80	100	88	92	44	36	44	56
DBLLTM-SK	M	75.9	73.1	78.2	75.3	76.6	74.2	71.6	76.3	79.8
DBLLTM-SK	S	17.6	16.4	18.2	16.3	18.8	15.6	17.4	17.0	19.2
DBLLTM-SK	N	180	80	100	88	92	44	36	44	56
PSTRNG-SK	M	85.0	81.2	88.1	81.4	88.4	78.8	84.1	84.1	91.2
PSTRNG-SK	S	22.3	23.7	20.8	25.8	17.9	26.3	20.0	25.3	16.0
PSTRNG-SK	N	180	80	100	88	92	44	36	44	56
PSTRNF-SK	M	49.5	42.0	55.4	44.0	54.7	38.1	46.8	49.9	59.8
PSTRNF-SK	S	29.5	28.4	29.1	31.4	26.7	29.5	26.5	32.4	25.7
PSTRNF-SK	N	180	80	100	88	92	44	36	44	56
VCFDCL-SK	M	67.5	68.0	67.1	54.5	79.8	57.9	80.3	51.2	79.4
VCFDCL-SK	S	19.5	20.2	19.0	18.7	9.9	22.0	6.6	14.3	11.6
VCFDCL-SK	N	181	80	101	88	93	44	36	44	57
DEFNTN-SK	M	36.3	41.2	32.8	18.5	66.3	28.7	72.8	8.9	63.3
DEFNTN-SK	S	36.0	38.1	34.1	29.0	24.8	34.4	27.9	18.9	23.1
DEFNTN-SK	N	126	53	73	79	47	38	15	41	32
CMPPRB-SK	M	29.7	31.6	28.4	12.9	57.9	16.1	65.6	8.1	54.3
CMPPRB-SK	S	36.3	37.6	35.5	25.1	34.9	32.1	27.9	14.9	37.6
CMPPRB-SK	N	126	53	73	79	47	38	15	41	32

3

Table 12

Stanford Foundation Skills Test - Spanish:
Descriptive Statistics for Each Scale for First-grade Bilingual Sample
Overall and by Language Category

Scale	Statistic	Overall	Low Eng	High Eng	Low Span	High Span	Lo Lo	Lo Hi	Hi Lo	Hi Hi
ALPHPR-SF	M	26.2	28.5	23.6	15.3	30.0	15.6	36.3	14.6	25.0
ALPHPR-SF	S	22.0	23.5	20.3	10.1	23.8	10.0	25.9	11.9	21.2
ALPHPR-SF	N	61	32	29	16	45	12	20	4	25
WRDNAM-SF	M									
WRDNAM-SF	S									
WRDNAM-SF	N									
SNGLTM-SF	M	96.3	95.3	97.4	93.6	97.3	92.9	96.7	95.7	97.7
SNGLTM-SF	S	5.3	6.2	3.9	7.6	3.9	8.3	4.2	5.5	3.7
SNGLTM-SF	N	62	32	30	16	46	12	20	4	26
DBLLTM-SF	M	91.7	89.7	93.8	91.5	91.8	91.2	88.9	92.4	94.0
DBLLTM-SF	S	9.3	10.3	7.6	7.3	9.9	8.3	11.5	4.1	8.0
DBLLTM-SF	N	62	32	30	16	46	12	20	4	26
PSTRNG-SF	M	92.0	89.9	94.3	83.3	95.0	79.2	96.2	95.8	94.0
PSTRNG-SF	S	18.1	21.5	13.6	28.6	11.6	32.2	6.1	5.3	14.5
PSTRNG-SF	N	62	32	30	16	46	12	20	4	26
PSTRNF-SF	M	72.4	67.1	78.1	63.8	75.4	59.1	71.9	77.8	78.1
PSTRNF-SF	S	24.8	25.6	22.9	29.6	22.4	32.6	19.7	11.7	24.4
PSTRNF-SF	N	62	32	30	16	46	12	20	4	26
VCFDCL-SF	M	82.5	81.6	83.4	70.3	86.7	69.7	88.8	72.2	85.2
VCFDCL-SF	S	17.9	21.3	13.7	27.2	10.8	31.0	6.3	13.2	13.2
VCFDCL-SF	N	62	32	30	16	46	12	20	4	26
DEFNTN-SF	M									
DEFNTN-SF	S									
DEFNTN-SF	N									
CMPPRB-SF	M									
CMPPRB-SF	S									
CMPPRB-SF	N									

Table 13

Stanford Foundation Skills Test - Spanish
Descriptive Statistics for Each Scale for Site 0 Bilingual Sample
Overall and by Language Category

Scale	Statistic	Overall	Lo Lo-0	Lo Hi-3	Hi Lo-0	Hi Hi-0
ALPHPR-GK	M	4.6	0.3	1.7		7.7
ALPHPR-GK	S	6.8	0.8	2.2		8.2
ALPHPR-GK	N	31	5	10		16
WRDMAN-GK	M					
WRDMAN-GK	S					
WRDMAN-GK	N					
SNGLTR-GK	M	91.2	90.9	88.9		92.7
SNGLTR-GK	S	6.6	5.5	7.9		6.1
SNGLTR-GK	N	31	5	10		16
DBLLTR-GK	M	77.8	58.9	79.7		82.5
DBLLTR-GK	S	17.3	20.4	14.4		14.7
DBLLTR-GK	N	31	5	10		16
PSTRND-GK	M	86.2	77.8	85.0		89.8
PSTRND-GK	S	10.2	7.9	11.0		8.9
PSTRND-GK	N	30	5	10		15
PSTRNF-GK	M	66.0	44.0	66.9		72.7
PSTRNF-GK	S	20.5	15.1	17.9		19.3
PSTRNF-GK	N	30	5	10		15
VCFBCL-GK	M	83.3	76.0	81.0		87.0
VCFBCL-GK	S	7.2	4.1	7.5		5.4
VCFBCL-GK	N	31	5	10		16
DEFNTN-GK	M					
DEFNTN-GK	S					
DEFNTN-GK	N					
CMPPRD-GK	M					
CMPPRD-GK	S					
CMPPRD-GK	N					
ALPHPR-GF	M	23.0	17.0	21.7		25.6
ALPHPR-GF	S	19.3	9.9	13.0		23.2
ALPHPR-GF	N	33	6	9		18
WRDMAN-GF	M					
WRDMAN-GF	S					
WRDMAN-GF	N					
SNGLTR-GF	M	97.4	96.7	96.0		97.9
SNGLTR-GF	S	4.6	5.8	5.0		4.1
SNGLTR-GF	N	34	6	9		19
DBLLTR-GF	M	93.9	93.4	91.7		94.4
DBLLTR-GF	S	9.3	5.2	12.2		9.1
DBLLTR-GF	N	34	6	9		19
PSTRND-GF	M	92.5	80.6	94.1		95.5
PSTRND-GF	S	13.8	27.7	7.2		7.4
PSTRND-GF	N	34	6	9		19
PSTRNF-GF	M	81.3	67.3	82.6		85.1
PSTRNF-GF	S	19.8	32.4	13.2		14.6
PSTRNF-GF	N	34	6	9		19
VCFBCL-GF	M	88.8	76.5	93.6		90.4
VCFBCL-GF	S	14.6	31.6	3.1		6.7
VCFBCL-GF	N	34	6	9		19
DEFNTN-GF	M					
DEFNTN-GF	S					
DEFNTN-GF	N					
CMPPRD-GF	M					
CMPPRD-GF	S					
CMPPRD-GF	N					

3.07

BEST COPY AVAILABLE

Table 14

Stanford Foundation Skills Test - Spanish:
Descriptive Statistics for Each Scale for Site 1 Bilingual Sample
Overall and by Language Category

Scale	Statistic	Overall	Lo Lo-1	Lo Hi-1	Hi Lo-1	Hi Hi-1
ALPHPR-SK	M	9.5		8.8	11.7	10.0
ALPHPR-SK	S	6.5		9.0	0.0	2.4
ALPHPR-SK	N	7		4	1	2
WRDNAM-SK	M					
WRDNAM-SK	S					
WRDNAM-SK	N					
SNDLTR-SK	M	86.2		61.5	94.3	97.1
SNDLTR-SK	S	10.4		9.5	0.0	0.0
SNDLTR-SK	N	6		4	1	1
BGLLTR-SK	M	79.2		70.2	94.4	100.0
BGLLTR-SK	S	19.2		16.9	0.0	0.0
BGLLTR-SK	K	6		4	1	1
PSTRND-SK	M	75.0		71.5	86.1	79.2
PSTRND-SK	S	23.7		32.2	0.0	23.5
PSTRND-SK	N	7		4	1	2
PSTRNF-SK	M	40.0		35.9	43.6	49.1
PSTRNF-SK	S	23.5		23.6	0.0	33.4
PSTRNF-SK	N	7		4	1	2
VCFBCL-SK	M	80.3		82.1	[illegible]	85.9
VCFBCL-SK	S	8.7		2.1	0.0	5.9
VCFBCL-SK	N	7		4	1	2
DEFNTN-SK	M					
DEFNTN-SK	S					
DEFNTN-SK	N					
CMPPRD-SK	M					
CMPPRD-SK	S					
CMPPRD-SK	N					
ALPHPR-SF	M	21.5	14.6	25.9	20.0	43.3
ALPHPR-SF	S	12.0	8.7	15.3	11.7	0.0
ALPHPR-SF	N	9	4	2	2	1
WRDNAM-SF	M					
WRDNAM-SF	S					
WRDNAM-SF	N					
SNDLTR-SF	M	91.4	85.0	97.2	94.3	100.0
SNDLTR-SF	S	8.7	8.2	4.0	8.1	0.0
SNDLTR-SF	N	9	4	2	2	1
BGLLTR-SF	M	92.0	90.3	90.3	93.1	100.0
BGLLTR-SF	S	7.3	9.5	5.9	5.9	0.0
BGLLTR-SF	N	9	4	2	2	1
PSTRND-SF	M	92.3	92.4	88.9	91.7	100.0
PSTRND-SF	S	6.9	9.4	4.0	3.9	0.0
PSTRND-SF	N	9	4	2	2	1
PSTRNF-SF	M	69.9	72.3	54.6	69.1	92.7
PSTRNF-SF	S	14.1	4.3	20.6	10.3	0.0
PSTRNF-SF	H	9	4	2	2	1
VCFBCL-SF	M	81.1	78.2	90.1	75.4	85.9
VCFBCL-SF	S	7.1	6.1	2.0	1.1	0.0
VCFBCL-SF	N	9	4	2	2	1
DEFNTN-SF	M					
DEFNTN-SF	S					
DEFNTN-SF	N					
CMPPRD-SF	M					
CMPPRD-SF	S					
CMPPRD-SF	N					

Table 15

Stanford Foundation Skills Test - Spanish
Descriptive Statistics for Each Scale for Site 2 Bilingual Sample
Overall and by Language Category

Scale	Statistic	Overall	Lo Lo-2	Lo Hi-2	Hi Lo-2	Hi Hi-2
ALPHPR-GK	M	7.0	1.7	5.0	5.9	10.6
ALPHPR-GK	S	5.5	0.0	3.9	1.2	6.6
ALPHPR-GK	N		1	7	2	6
WRDMAN-GK	M					
WRDMAN-GK	S					
WRDMAN-GK	N					
SNDLTR-GK	M	89.7	91.4	85.7	77.2	97.1
SNDLTR-GK	S	11.0	0.0	9.9	20.2	4.4
SNDLTR-GK	N	17	1	7	2	7
BGLLTR-GK	M	82.4	72.2	69.1	88.9	95.2
BGLLTR-GK	S	16.6	0.0	16.0	7.8	5.9
BGLLTR-GK	N	17	1	7	2	7
PSTRNG-GK	M	76.6	77.8	86.9	44.5	75.4
PSTRNG-GK	S	28.6	0.0	20.7	15.8	35.2
PSTRNG-GK	N	17	1	7	2	7
PSTRNF-GK	M	33.6	14.5	33.1	3.7	43.4
PSTRNF-GK	S	28.0	0.0	22.6	5.2	33.1
PSTRNF-GK	N	17	1	7	2	7
VCFBCL-GK	M	77.5	73.2	80.1	69.7	77.7
VCFBCL-GK	S	8.1	0.0	7.7	1.0	9.1
VCFBCL-GK	N	17	1	7	2	7
DEFNTN-GK	M					
DEFNTN-GK	S					
DEFNTN-GK	N					
CMPPRO-GK	M					
CMPPRO-GK	S					
CMPPRO-GK	N					
ALPHPR-GF	M	33.9	13.4	53.3	9.2	20.0
ALPHPR-GF	S	28.5	18.9	28.0	12.9	15.3
ALPHPR-GF	N	19	2	9	2	6
WRDMAN-GF	M					
WRDMAN-GF	S					
WRDMAN-GF	N					
SNDLTR-GF	M	96.7	97.2	96.5	97.2	96.7
SNDLTR-GF	S	3.1	4.0	3.7	4.0	2.1
SNDLTR-GF	N	19	2	9	2	6
BGLLTR-GF	M	87.7	80.6	85.8	91.7	[illegible]
BGLLTR-GF	S	9.0	3.9	11.8	3.9	3.0
BGLLTR-GF	N	19	2	9	2	6
PSTRNG-GF	M	90.9	48.6	100.0	100.0	88.4
PSTRNG-GF	S	27.1	68.7	0.0	0.0	28.3
PSTRNG-GF	N	19	2	9	2	6
PSTRNF-GF	M	57.7	8.2	65.0	86.4	53.6
PSTRNF-GF	S	30.8	11.6	21.0	1.3	35.4
PSTRNF-GF	N	19	2	9	2	6
VCFBCL-GF	M	72.0	32.4	83.7	69.0	68.5
VCFBCL-GF	S	21.9	45.8	5.4	21.9	16.6
VCFBCL-GF	N	19	2	9	2	6
DEFNTN-GF	M					
DEFNTN-GF	S					
DEFNTN-GF	N					
CMPPRO-GF	M					
CMPPRO-GF	S					
CMPPRO-GF	N					

Table 16

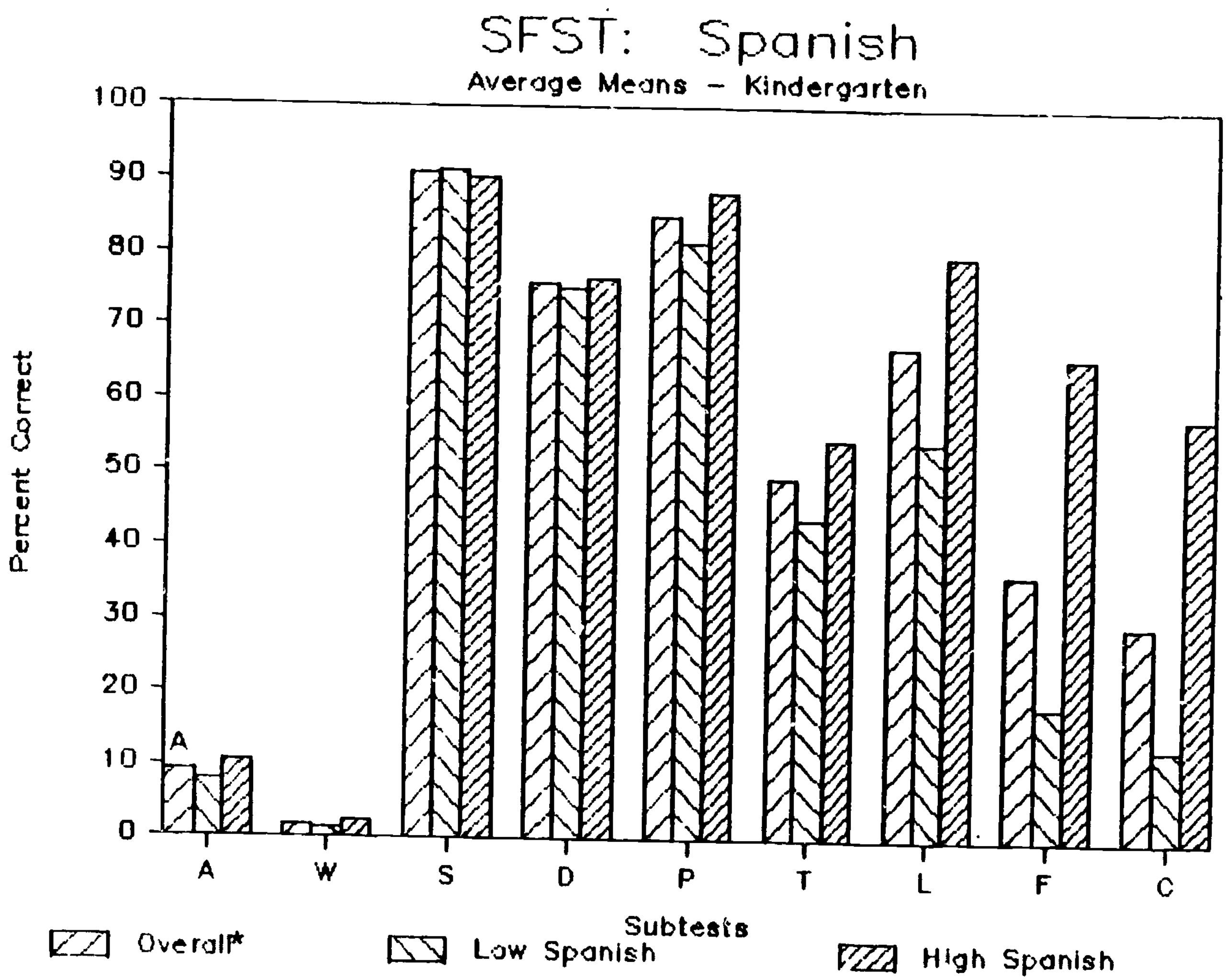
Stanford Foundation Skills Test - Spanish:
Descriptive Statistics for Each Scale for Site 3 Bilingual Sample
Overall and by Language Category

Scale	Statistic	Overall	Lo Lo-3	Lo Hi-3	Hi Lo-3	Hi Hi-3
ALPHPR-SK	M	7.3	6.8		7.0	10.2
ALPHPR-SK	S	6.7	5.2		6.9	10.0
ALPHPR SK	N	68	25		35	8
WRDNAM-SK	M	1.8	0.7		1.7	6.3
WRDNAM-SK	S	5.0	2.3		4.9	8.6
WRDNAM-SK	N	68	25		35	8
SNGLTM-SK	M	91.8	88.8		94.2	90.4
SNGLTM-SK	S	6.5	6.8		5.5	5.9
SNGLTM-SK	N	69	25		36	8
DBLLTM-SK	M	75.9	74.1		77.9	72.6
DBLLTM-SK	S	15.7	14.7		16.0	18.3
DBLLTM-SK	N	69	25		36	8
PSTRNG-SK	M	81.4	70.0		85.6	98.3
PSTRNG-SK	S	27.4	31.1		24.8	3.3
PSTRNG-SK	N	69	25		36	8
PSTRNF-SK	M	44.3	31.9		51.4	51.6
PSTRNF-SK	S	32.2	32.7		31.5	23.8
PSTRNF-SK	N	69	25		36	8
VCFDCL-SK	M	49.8	45.2		48.4	70.9
VCFDCL-SK	S	16.2	17.7		12.5	10.2
VCFDCL-SK	N	69	25		36	8
DEFNTN-SK	M	13.9	14.3		5.1	52.1
DEFNTN-SK	S	23.6	26.4		9.0	23.5
DEFNTN-SK	N	69	25		36	8
CMPPRB-SK	M	8.2	3.6		5.7	34.1
CMPPRB-SK	S	15.4	6.4		9.7	28.5
CMPPRB-SK	N	69	25		36	8

Table 17

Stanford Foundation Skills Test - Spanish:
Descriptive Statistics for Each Scale for Site 5 Bilingual Sample
Overall and by Language Category

Scale	Statistic	Overall	Lo Lo-5	Lo Hi-5	Hi Lo-5	Hi Hi-5
ALPHPR-SK	M	14.8	9.2	10.4	24.7	18.6
ALPHPR-SK	S	20.1	23.4	8.8	29.1	21.0
ALPHPR-SK	N	57	13	15	5	24
WRDNAM-SK	M	1.9	2.6	1.7	1.7	1.7
WRDNAM-SK	S	5.5	9.2	4.7	3.7	3.4
WRDNAM-SK	N	57	13	15	5	24
SNGLTM-SK	M	90.5	92.1	88.9	88.0	91.1
SNGLTM-SK	S	8.2	4.7	8.8	9.6	9.2
SNGLTM-SK	N	57	13	15	5	24
DBLLTM-SK	M	72.7	80.6	67.8	55.6	75.0
DBLLTM-SK	S	19.8	12.8	19.7	11.6	22.1
DBLLTM-SK	N	57	13	15	5	24
PSTRNG-SK	M	92.3	96.1	85.6	88.9	95.3
PSTRNG-SK	S	14.9	6.3	21.5	24.9	9.1
PSTRNG-SK	N	57	13	15	5	24
PSTRNF-SK	M	52.8	49.8	41.8	58.6	60.0
PSTRNF-SK	S	27.2	24.7	27.9	36.8	25.0
PSTRNF-SK	N	57	13	15	5	24
VCFDCL-SK	M	75.8	74.2	79.4	61.7	77.2
VCFDCL-SK	S	14.3	17.9	6.6	20.2	13.4
VCFDCL-SK	N	57	13	15	5	24
DEFNTN-SK	M	63.4	56.4	72.8	36.7	67.0
DEFNTN-SK	S	29.0	31.6	27.9	41.9	22.2
DEFNTN-SK	N	57	13	15	5	24
CMPPRB-SK	M	55.7	46.1	65.6	25.6	61.1
CMPPRB-SK	S	37.3	42.6	27.9	30.8	38.3
CMPPRB-SK	N	57	13	15	5	24



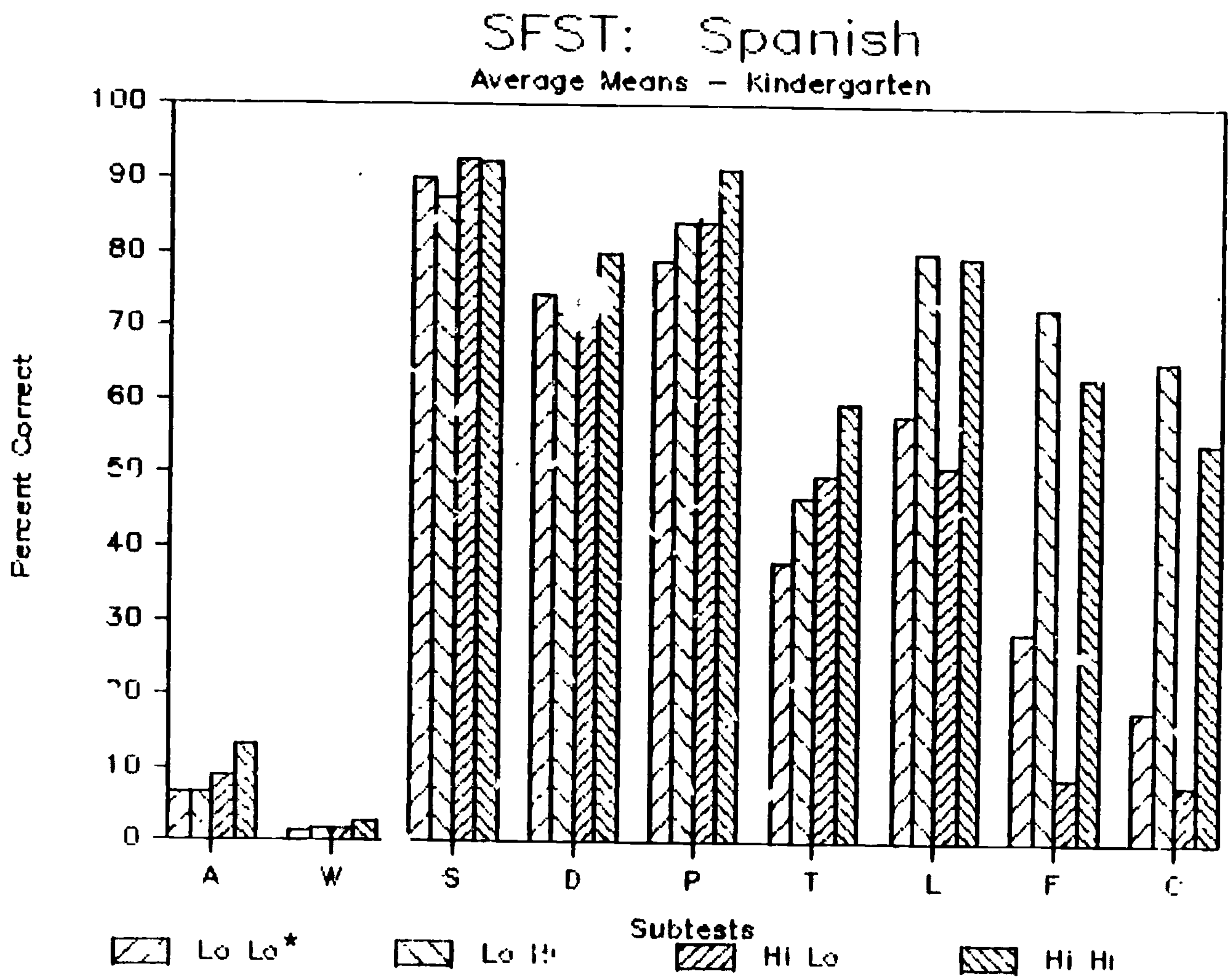
A = Alphabet Knowledge
W = Word Naming
S = Single-Letter Matching

D = Double-Letter Matching
P = Auditory-Phonetic Segmentation: Training
T = Auditory-Phonetic Segmentation: Transfer

L = Fine Distinctions Common Labels
F = Definitions
C = Comprehension

*N≈126 for the W, F, and C scales; otherwise N≈180.

Figure 6. SFST-S scale means for the kinderg -ten entry bilingual sample overall and by Spanish entry category.



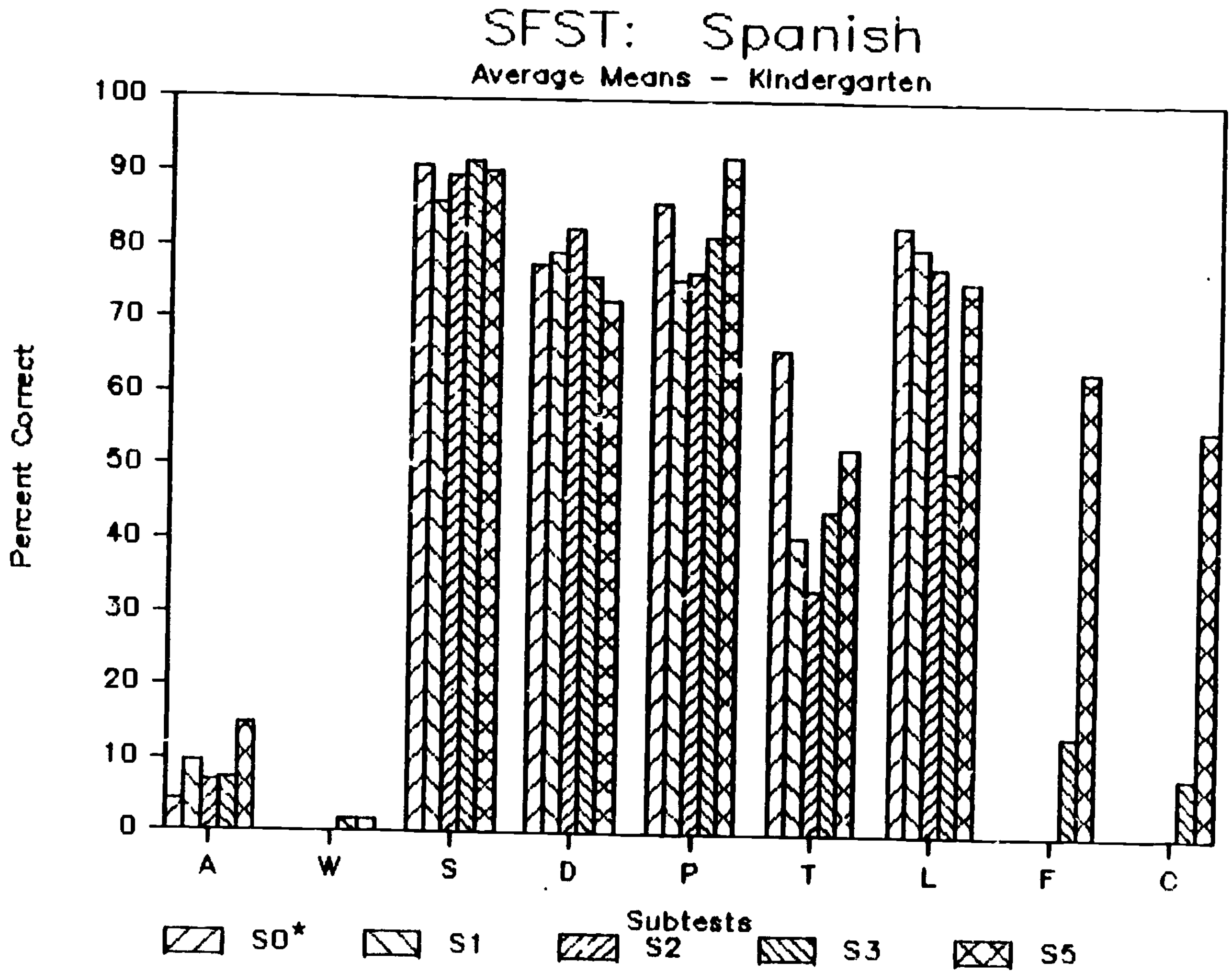
A = Alphabet Knowledge
W = Word Naming
S = Single-Letter Matching

D = Double-Letter Matching
P = Auditory-Phonetic Segmentation: Training
T = Auditory-Phonetic Segmentation: Transfer

L = Fine Distinctions Common Labels
F = Definitions
C = Comprehension

*Low English, Low Spanish entry.

Figure 7. SFST-S scale means for the kindergarten entry bilingual sample by English-Spanish entry category.



A = Alphabet Knowledge
W = Word Naming
S = Single-Letter Matching

D = Double-Letter Matching
P = Auditory-Phonetic Segmentation: Training
T = Auditory-Phonetic Segmentation: Transfer

L = Fine Distinctions
Common Labels
F = Definitions
C = Comprehension

*Site 0.

Figure 8. SFST-S scale means for the kindergarten entry bilingual sample by site.

both Training and Transfer tasks, and (b) for Definitions and Comprehension, skill at Site 5 is superior to Site 3, suggesting the greater Spanish skills of these students.

In examining the tabled data for each site, the general pattern of the high Spanish groups outperforming the low Spanish groups appears to hold.

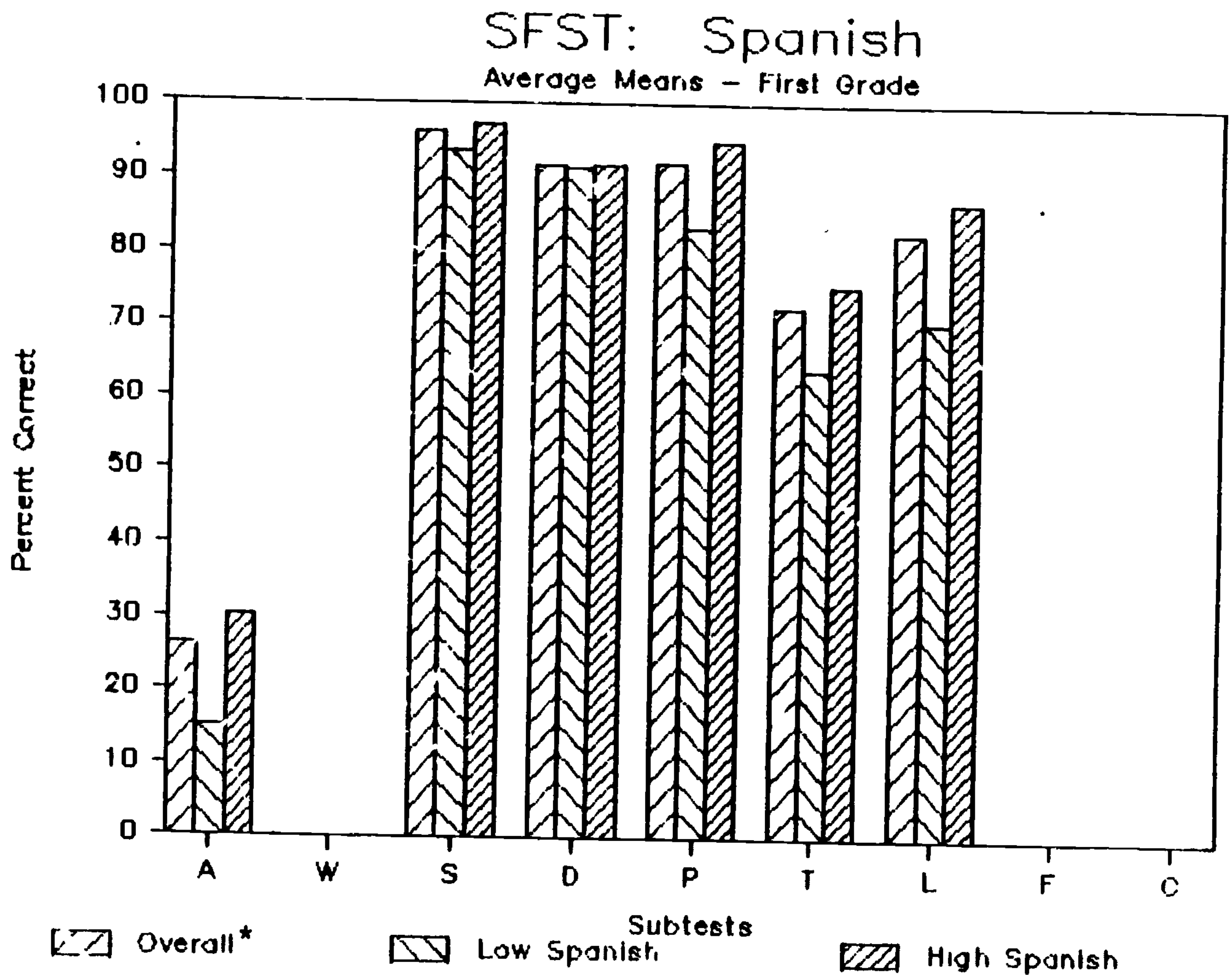
Figures 9 and 10 (based on Table 12) present the first-grade entry data obtained for those border site students who entered the study as first-graders. Again, performance follows that described for the individual cohorts, and, in general, is higher than that obtained from the kindergarten sample. Interestingly, there does not seem to be any significant difference between the four language entry categories, except that the group low in both English and Spanish tends to show the lowest level of skill for each of the tasks, most notable in those of Auditory-Phonetic Segmentation.

Correlations

The correlations, both within and between the English and Spanish administrations for each of the two grade levels, are presented in the four panels of Table 18. The first and second panels present the within language correlations for the kindergarten and first-grade administrations, respectively. In these two panels, the English correlations are presented above the diagonal, with the Spanish correlations below. In the bottom two panels, the between language correlations are presented, with the rows representing the English scales and the columns representing the Spanish scales. The kindergarten coefficients are presented first, then those for first-grade. In discussing these matrices, again, one must keep in mind the structure of the sample given in Table 1. In the material which follows, only correlations which show a minimum of 15% shared variance (significant at the .001 level based on the lowest sample size) will be discussed.

First, consideration will be given to the correlations between scale summary measures for the English administrations at kindergarten (first panel, above the diagonal). The sample size for each of these coefficients is approximately 162, except for the tasks of Word Naming, Definitions, and Comprehension, which is 122, and represents only the non-border sites. The pattern of correlations is easily summarized: (a) the non-linguistic tasks of Single- and Double-Letter Matching are related, (b) the linguistic tasks of Vocabulary, Definitions, and Comprehension are all inter-related, (c) the two Auditory-Phonetic Segmentation tasks, and the Vocabulary task are inter-related, and (d) Alphabet Knowledge is related to all tasks except Letter-Matching and the Training task in Auditory-Phonetic Segmentation.

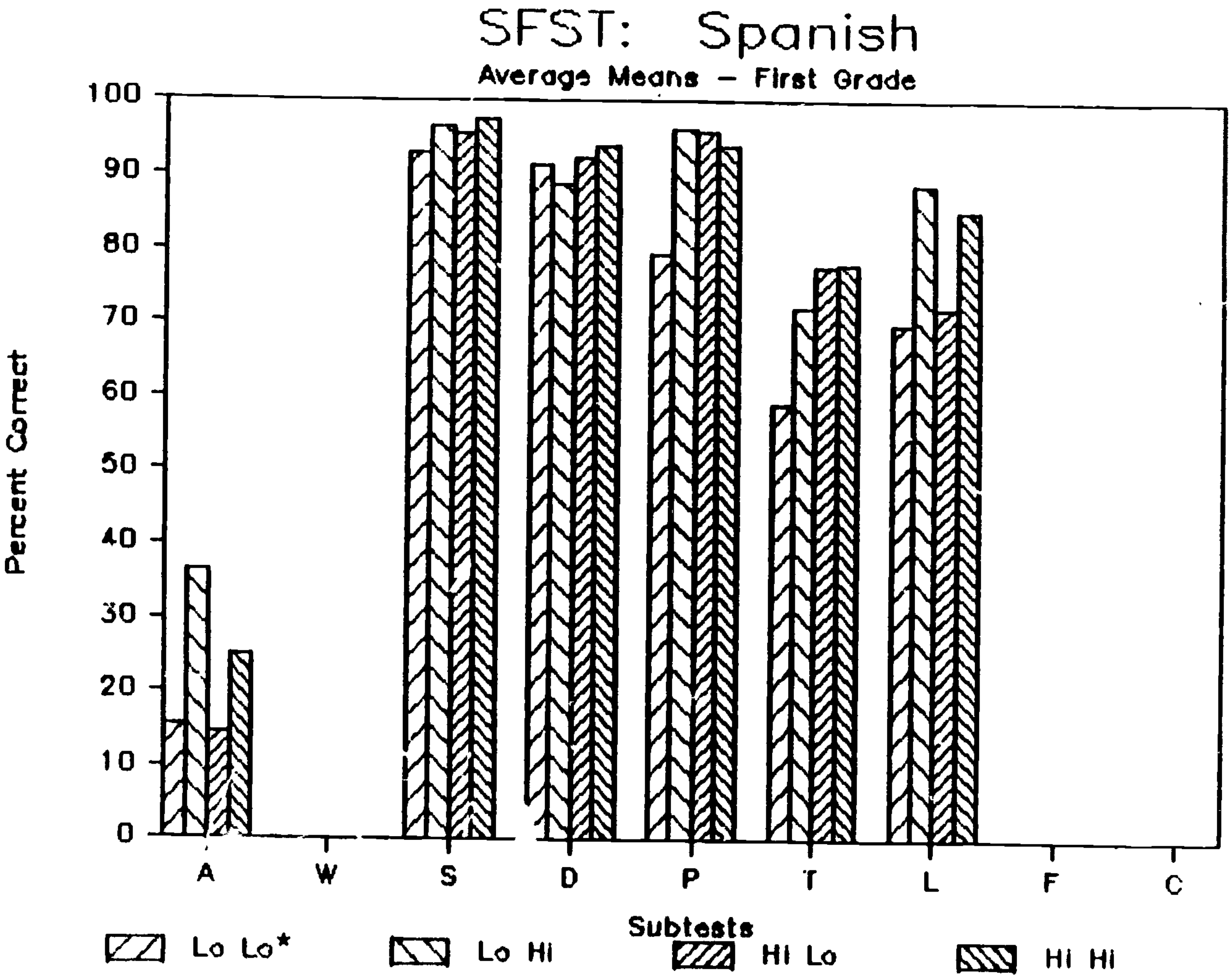
Keeping in mind that the highest correlation coefficient within this set is .56, the following generalizations are appropriate: (a) there appears to be some general language measure which is tapped by knowledge of the alphabet, (b) skill in visual matching is fairly independent of the other pre-reading skills assessed, (c) there is a



A = Alphabet Knowledge
W = Word Naming
S = Single-Letter Matching
D = Double-Letter Matching
P = Auditory-Phonetic Segmentation: Training
T = Auditory-Phonetic Segmentation: Transfer
L = Fine Distinctions Common Labels
F = Definitions
C = Comprehension

*N=62.

Figure 9. SFST-S scale means for the first-grade entry bilingual sample overall and by Spanish entry category.



A = Alphabet Knowledge
W = Word Naming
S = Single-Letter Matching

D = Double-Letter Matching
P = Auditory-Phonetic Segmentation: Training
T = Auditory-Phonetic Segmentation: Transfer

L = Fine Distinctions Common Labels
F = Definitions
C = Comprehension

*Low English, Low Spanish entry.

Figure 10. SFST-S scale means for the first-grade entry bilingual sample by English-Spanish entry category.

Stanford Foundation Skills Test:
Correlations within and Between English and Spanish Administrations
for Kindergarten and First Grade Bilingual Sample

SFST - English Kinder

		ALPHPR	WRDNAM	SNGLTM	DBLLTM	PSTRNG	PSTRNF	VCFDCL	DEFNTN	CMPPRB
SFST Spanish Kinder	ALPHPR	-	0.53	0.29	0.30	0.30	0.43	0.51	0.44	0.37
	WRDNAM	0.46	-	0.21	0.25	0.10	0.22	0.27	0.28	0.23
	SNGLTM	0.11	0.14	-	0.54	0.27	0.30	0.18	0.19	0.22
	DBLLTM	0.17	0.11	0.58	-	0.31	0.33	0.29	0.16	0.19
	PSTRNG	0.18	0.08	0.21	0.17	-	0.51	0.51	0.22	0.26
	PSTRNF	0.21	0.16	0.23	0.29	0.57	-	0.48	0.38	0.36
	VCFDCL	0.18	0.17	.00	0.07	0.24	0.22	-	0.56	0.56
	DEFNTN	0.34	0.16	0.01	0.01	0.25	0.15	0.80	-	0.52
	CMPPRB	0.31	0.11	0.07	0.06	0.24	0.19	0.75	0.76	-

SFST - English First

		ALPHPR	WRDNAM	SNGLTM	DBLLTM	PSTRNG	PSTRNF	VCFDCL	DEFNTN	CMPPRB
SFST Spanish First	ALPHPR	-	-	0.16	0.14	0.30	0.21	0.68	-	-
	WRDNAM	-	-	-	-	-	-	-	-	-
	SNGLTM	0.09	-	-	0.45	0.03	0.38	0.31	-	-
	DBLLTM	0.07	-	0.29	-	-0.14	0.57	0.25	-	-
	PSTRNG	0.20	-	-0.09	0.02	-	0.28	0.26	-	-
	PSTRNF	0.11	-	-0.04	0.25	0.60	-	0.40	-	-
	VCFDCL	0.28	-	0.09	0.12	0.65	0.62	-	-	-
	DEFNTN	-	-	-	-	-	-	-	-	-
	CMPPRB	-	-	-	-	-	-	-	-	-

SFST - Spanish Kinder

		ALPHPR	WRDNAM	SNGLTM	DBLLTM	PSTRNG	PSTRNF	VCFDCL	DEFNTN	CMPPRB
SFST English Kinder	ALPHPR	0.57	0.38	0.30	0.32	0.19	0.35	.00	0.10	0.10
	WRDNAM	0.24	0.53	0.27	0.29	0.15	0.21	0.02	-0.05	.00
	SNGLTM	0.14	0.14	0.48	0.38	0.16	0.13	0.19	0.20	0.17
	DBLLTM	0.17	0.07	0.48	0.67	0.21	0.15	0.08	0.11	0.09
	PSTRNG	0.20	0.07	0.36	0.21	0.34	0.20	-0.09	0.08	0.11
	PSTRNF	0.15	0.10	0.30	0.24	0.40	0.59	-0.05	0.03	0.09
	VCFDCL	0.20	0.14	0.23	0.19	0.06	0.15	-0.29	-0.20	-0.09
	DEFNTN	0.31	0.16	0.22	0.18	0.26	0.28	0.17	0.13	0.13
	CMPPRB	0.31	0.04	0.23	0.26	0.27	0.34	0.03	0.00	0.12

SFST - Spanish First

		ALPHPR	WRDNAM	SNGLTM	DBLLTM	PSTRNG	PSTRNF	VCFDCL	DEFNTN	CMPPRB
SFST English First	ALPHPR	0.14	-	0.25	0.16	-0.02	0.01	-0.11	-	-
	WRDNAM	-	-	-	-	-	-	-	-	-
	SNGLTM	0.01	-	0.53	0.33	-0.07	0.23	0.22	-	-
	DBLLTM	-0.33	-	0.26	0.60	0.03	0.54	0.23	-	-
	PSTRNG	0.25	-	-0.01	-0.23	0.39	0.18	0.24	-	-
	PSTRNF	-0.17	-	0.12	0.26	0.14	0.63	0.25	-	-
	VCFDCL	0.04	-	0.20	0.19	0.65	0.16	0.05	-	-
	DEFNTN	-	-	-	-	-	-	-	-	-
	CMPPRB	-	-	-	-	-	-	-	-	-

dimension of word knowledge which is evidenced in skill in isolated word tasks and in comprehension of connected text, and (d) the metalinguistic task of phonetic segmentation is relatively independent of the other measures, but its relation with the Vocabulary task suggests that knowledge of the language aids in both the acquisition and transfer of the skill as it is assessed in this test.

Next consideration is given to the between scale correlations for the Spanish administrations at kindergarten (Table 18, first panel, below the diagonal). The sample size for each of these coefficients is approximately 180, except for the tasks of Word Naming, Definitions, and Comprehension, which is 126, and again, represents only the non-border sites. With two exceptions, the pattern of correlations here is quite similar to those just described: (a) the non-linguistic tasks of Single- and Double-Letter Matching are related, (b) the linguistic tasks of Vocabulary, Definitions, and Comprehension are all inter-related, (c) the two Auditory-Phonetic Segmentation tasks are related, but unlike the English version, these show little relation to the Vocabulary task, and (d) Alphabet Knowledge is only related to sight-word recognition skill, unlike the broader base of relations it maintained in English.

In the Spanish kindergarten data, the correlation coefficients discussed above are generally higher than those found in English, the greatest being .80. With this in mind, the following generalizations are warranted: (a) alphabet knowledge in Spanish does not seem indicative of some general language skill as it seems to be in English, (b) skills in both visual matching and phonetic segmentation are independent of the other pre-reading skills assessed (to a greater degree here than in English), and (c) there is a dimension of word knowledge which is evidenced in skill in isolated word tasks and in comprehension of connected text.

Before discussing the first-grade data, the next focus will be on the relations between the English and Spanish scales based on the kindergarten sample, which are presented in the third panel of Table 18. For the Definitions and Comprehension tasks, given only in the non-border sites, no correlation reaches the 15% shared variance criterion. For the remaining seven scales, the largest correlations fall on the diagonal, suggesting that for any scale in one language, its strongest relationship within the other language test is with its corresponding scale. Of these seven diagonal coefficients, all are positive, except for the Vocabulary coefficient, which is negative, and shares less than 15% of the variance between the language measures. The other coefficients are above this value (ranging from .48 to .66), except for training in the phonetic segmentation task.

Thus, this pattern of results suggests that (a) the visual matching tasks and the metalinguistic task of phonetic segmentation possess a degree of transferability between the two languages, (b) the linguistic tasks of Vocabulary, Definitions, and Comprehension are each independent <u>across</u> (but not within) the two languages, and (c) alphabet knowledge and sight-word recognition tend to be related across the two languages. The independence of the linguistic tasks would be expected, as would the non-independence of the non-linguistic tasks. However, the suggested transferability of the metalinguistic task which is thought by many (Gough & Hillinger, 1979; Tunmer, Pratt & Herriman, 1984) to be of critical import in the acquisition of reading skill in any alphabetic reading system, is not necessarily expected, and has significant implications for bilingual instruction. Further, at least for the literacy skills that these English-Spanish bilingual students bring to school, the suggestion that sight-word recognition (or rudimentary decoding skill) in one language is related to that in a second, is likewise significant.

Turning now to the first-grade administrations displayed in the second panel of Table 18. Here, between scale correlations for the English administration are given above the diagonal, Spanish below the diagonal. Again, recall that the sample is small (50 and 62 students in the English and Spanish administrations, respectively), comes entirely from the border sites, and does not contain data for the Word Naming, Definitions, or Comprehension components. Nonetheless, the pattern of results is quite similar. For English, (a) knowledge of the alphabet is related to vocabulary skill, (b) the two visual matching tasks are related, but fairly independent of the other tasks, and (c) for phonetic segmentation, while the training skill is not related to any of the other scales, the transfer task shows the same relationship to vocabulary skill (but also, shows a novel relationship with Double-Letter Matching). For Spanish, as in the kindergarten data, (a) no relationship for knowledge of the alphabet obtains, (b) the correlations for the visual matching tasks, while all are below the 15% criterion, are, nonetheless, highest between each other, and (c) the phonetic segmentation tasks and the vocabulary task are all interrelated (at about .62).

The pattern of between language correlations for the first-grade data (bottom panel of Table 18) are, again, similar to those found at kindergarten: the largest correlations fall along the diagonal (with the exception of the puzzling relationship between English Double-Letter Matching and Spanish Auditory-Phonetic Segmentation Transfer). Thus, these data support the arguments given above in the discussion of the kindergarten sample.

Reading Achievement Measures

As stated above, growth in reading was assessed through multiple information sources: the Interactive Reading Assessment System, standardized test scores, and Informal Reading Inventories. The first two of these are discussed below, providing details of the tasks, materials, scoring, reliability, and descriptive statistics on the performance of the bilingual sample.

Interactive Reading Assessment System

The Interactive Reading Assessment System - IRAS was employed for assessing the student's reading ability. The IRAS, an individually administered diagnostic assessment system, was designed for research application initially by Dr. Robert Calfee and his associates at Stanford University in 1974, and has undergone two major revisions (Calfee & Calfee, 1979; 1981). Modeled after the informal reading inventory, the IRAS provides independent measures of several component skills essential for fluent reading. The materials in the test were selected to cover a wide range of skills and knowledge in the areas of reading and oral language from the level usually expected of a mid-year first-grader to that of a junior high school student. The Spanish version of the IRAS was developed in 1979 using the same format and procedures as those used in the development of the English edition (Calfee, Calfee & Peña, 1979), and was updated as modifications were made in the English version.

The IRAS has undergone four revisions during its use in this study over the five-year data collection period (only the Year 3 and Year 4 instruments were identical). Most updates consisted of adding more difficult levels within a given task in order to keep pace with student skills as they grew from kindergarten through fourth grade, although some included new tasks (e.g., sentence reading and comprehension of expository texts in the third year) and/or changes in procedures.

Each test was individually administered by trained personnel, taking approximately 45 minutes to complete. The entire testing session was tape recorded and the tapes and individual protocols completed by the tester were given to trained in-house staff for scoring. All testing was done in the Spring (March to May), and all targets were tested beginning in first grade, and continuing until exit from the study. All targets in the bilingual sample were tested with both language versions. For these students, order of test administration was counterbalanced within each site, with approximately three weeks between testing. The monolingual English and Spanish target students were only tested in their respective language.

The specific structure for each of the IRAS subtests, with details of any incorporated modifications, and the scoring procedures employed, are presented below. The current English version of the system will be the primary focus of discussion, but the Spanish version was treated in an identical manner (except where noted).

Tasks, Materials, and Scoring

The rationale for the tasks appearing in the IRAS is based on a theory of reading as a set of independent component skills (Calfee & Drum, 1979). The areas of knowledge assessed in the system include: reading of isolated words, definition of common words within and beyond the student's reading vocabulary, and selected word analysis skills based on the pronunciation of synthetic words. Comprehension of connected text is also assessed, and in several contexts: reading and listening comprehension of both narrative texts (typical of those found in reading texts and literature series) and more difficult, expository texts.

Materials within each subtest are ordered by difficulty based upon grade-level expectations of performance, with each IRAS level roughly corresponding to a half-grade level. Thus, material contained within the fourth level of a given subtest corresponds to material which average second grade students should be able to handle. As mentioned above, the IRAS contains six separate subtests. Each one is discussed below, following the order in which it appeared in the testing procedure.

Real Word Decoding. The first assessment was that of the student's ability to decode real words. The materials consisted of 14 ordered six-word lists, where higher ordered lists contained increasingly more difficult words based on word frequency, number of syllables, and complexity of letter-sound correspondence. Words within a list were equated across these dimensions. The student was first presented with the lists and asked to indicate the most difficult list he thought he could successfully read. The student was then asked to read the selected list aloud.

If the student failed to read half or more of the words correctly, less difficult lists were presented until the student could read at least half the words in a presented list. Since the lists were ordered in terms of difficulty, higher order lists were not presented under the assumption that the student would not succeed on these more difficult items. Once a list was successfully passed, success was assumed for the less difficult, lower ordered lists based on the same rationale.

For a student who was successful on the first list attempted, more difficult lists were presented until the student failed to read correctly at least half of the words in a given list. Again, success was assumed for lists not presented which were of a lower order than those on which the student was able to successfully meet criterion, and failure was assumed for those lists not presented which were of a higher order than those on which the student failed to meet criterion.

In scoring the Real Word Decoding scale, each item was assigned a numeric value depending on the quality of the response. A value of 3 was assigned to items given completely correct (disregarding dialectical variations), a 2 for items which were mostly correct (e.g., completely correct except for a single consonant cluster or vowel), a 1 for items which were only partially correct (e.g., initial segment correct, but remaining segments incorrect), and a 0 for cases of wild or no attempted response.

In deriving a scale score, a "critical index" was computed based entirely on the lists attempted by the student. This index supplied information about which list was the highest ordered list on which the student succeeded, (thus allowing a comparison to grade level expectations), and also, the relative quality of performance within this highest success list. The index consisted of an integer value corresponding to the order of the list of highest success (ranging from 0 to 14), plus a decimal value which was the ratio of assigned points to total possible points on the list of highest success (a maximum of

18 points per list, derived from three points per item for each of the six items).

Thus, the integer portion of the critical index corresponds to the ordinal value of the most difficult list the student could read -- the higher the value, the more difficult the material the student could successfully complete. Concerning the decimal portions, low values represent poor, but passing, performance on the list of highest success (given that success in this task is defined as three completely correct responses, the lowest value possible for any successfully read list is .5 -- students failing to read any list successfully will have integer values of 0, and may have decimal values lower than .5). Relatively large decimal values correspond to high performance on the list of highest success (values of 1.0 assigned for perfect performance on the list of highest success were converted to .99). Thus, for this task, scale scores are bounded by 0.0 and 14.99. The computation of such critical indices was used throughout the IRAS, with the exception of Synthetic Word Spelling and Sentence Reading.

Over the five-year data collection period, the only modification made in this task for the English version occurred between the Year 2 and Year 3 administrations, and involved replacing a few of the words, mostly at the lower levels of the test, in order to obtain a slightly broader range of spelling patterns. In each case, the replacement words were of the same word frequency class as those which were removed. No changes were made in the Spanish version of this task over the entire data collection period.

Vocabulary Definitions. The next task contained 14 three-word lists, each list being a subset of the corresponding list of words used in the Real Word Decoding task discussed above. The student began this task with the lowest level list on which failure was obtained in Real Word Decoding. For each word read by the tester, the student was asked to define it. If an inadequate or questionable definition was given, the student was then asked if he could think of another word which meant the same thing. If this probe did not produce an adequate response, then the student was read three alternative definitions and asked to select the best one. The student was considered to be successful on a given list if he could produce an adequate response under any of the above conditions for at least two of the items. As in the previous task, the student was moved through the lists until that point was found where success was obtained on list n, but failure on list n + 1. Again, success was assumed for any untested lists below this point, and failure was assumed for any untested lists above it.

For the Vocabulary Definition task, each item was assigned a value ranging from 0 to 3 dependent upon the quality of the response. A value of 3 was assigned to any item for which the student gave either an adequate "dictionary definition," a fairly extensive functional definition, or a synonym. A value of 2 was given to poor, but acceptable definitions, associations, or unelaborated functional definitions. A value of 1 was given for correct multiple choice definitions if the responses to the first probes were inadequate, but the proper definition was selected among the three alternatives. A value of 0 was assigned whenever the student gave either a wild or no response to the first probes, and then made an incorrect selection in the multiple choice condition.

As in the Real Word Decoding task, a critical index was computed to characterize performance on the Definition task. Again, the integer portion of this value represented the order of the list of highest success, and the decimal part, the proportion of assigned points to total possible points (a maximum of 3 x 3 = 9) on this list of highest success. Thus, these scale values are bounded by 0.0 and 14.99.

As in the Real Word Decoding task, the Definition task was modified between the Year 2 and Year 3 administrations. During the first two years of testing, four items from each of the Real Word Decoding six-word lists were included in the Vocabulary Definition task, and the criterion for success on a given list was three correct responses from the four items. However, beginning with the Year 3 testing, the number of words the student was asked to define was reduced from four to three per list (and the success criterion was reduced from three to two) in an effort to reduce testing time. This change was employed in both the English and Spanish versions of the test. For the English version, however, some changes were also made in the lists of words to be defined in conjunction with the changes described above in the Real Word Decoding materials.

For students who did not have any success in the Real Word Decoding task, the following four tasks described, all requiring decoding skills (Synthetic Word Decoding, Synthetic Word Spelling, Sentence Reading, and Reading Comprehension), were not administered (assuming failure). For such students, the Alphabet Recognition task (not described in this report) was administered, followed by Listening Comprehension assessments (described after these four procedures).

Synthetic Word Decoding. In the third component of the IRAS, the student was presented with six lists of synthetic words, with the first four lists containing six items each and the remaining two lists containing nine items each. The synthetic words were constructed to correspond to the orthography of the language assessed, and lists were ordered by difficulty ranging from simple consonant-vowel-consonant patterns to blends, digraphs, vowel variations and polysyllabic items (e.g., for the English version, from hin and pame in the lowest ordered list to rhosmic and conspartable in the highest ordered list). Before being asked to read the synthetic words aloud, the student was told that he items were not real words, and they had no meaning, but that they could be pronounced like real words. Each student began this task on the easiest list, and proceeded to more difficult lists as long as responses were attempted on at least half of the items within a list.

For the lists of synthetic words, each of the items were scored as follows. A value of 3 was assigned to any item that was pronounced without error. A value of 2 was given to those responses that were mostly correct (e.g., correct responses except for a minor letter-sound error such as a vowel shift within vowel family, a stress variant, or pronunciation of a final "e"). A value of 1 was given for responses that were partly correct (e.g., correct responses except for a single vowel or consonant substitution or deletion). A value of 0 was assigned for assumed failure, no response, or for mispronunciations beyond those tolerated in the above categories. Note that the scoring was fairly stringent -- two major errors within an item were sufficient to receive a score of 0 (e.g., pronouncing affremiation as affrematon).

After examining the patterns of item difficulty over the 42 items, the order of presentation of the lists was modified for purposes of computing a critical index. Since almost all subjects who attempted the first word presented, responded to the entire set of synthetic words, the potential problem of assumed responses in such a re-ordering did not arise. The scoring order of the lists is presented below:

Scoring Order	Ordinal List Level
List 1 (items 1-6)	1
List 2 (items 1-6),List 3 (items 1-6)	2
List 5 (items 1-6)	3
List 4 (items 1-6)	4
List 6 (items 1, 3-6)	5
List 5 (items 7-9),List 6 (items 7-9)	6

Note that one item was deleted from List 6 as it yielded particularly idiosyncratic responses.

A critical index over these six lists was computed to represent student performance on this task. The integer portion represented the ordinal value of the list of highest success (ranging from 0 to 6), where success on a given list was achieved if at least half of the items in the list received a value of 1 or more. To this value a decimal was added which was the ratio of assigned points to total possible points on the list of highest success. Thus, scores for thi scale were bounded by 0.0 and 6.99. No attempt was made to tie the materials used in this task to those employed in Real Word Decoding, and thus, while the summary index for the latter can be related to grade level expectations of performance, this cannot be done for Synthetic Word Decoding.

The scoring procedure just described applied to the English IRAS administrations of Years 3 through 5, which were identical with respect to Synthetic Word Decoding. However, the Year 1 and 2 materials, which matched each other, were completely different from this later set, consisting of six six-word lists. Therefore, these materials were matched to those from the later administrations based on syllabic structure, letter-sound correspondence, number of letters, and number of phonemes. Based on these comparisons, a mapping of the earlier administration material into the first five lists of the later administration materials was affected (with one item being deleted from the set of 36). Thus, in Years 1 and 2, the computed critical index was bounded by 0.0 and 5.99.

For the Spanish IRAS, materials consisted of six six-word lists, which were not modified during the five-year data collection phase. The 36 items contained in this task were also re-organized based on analyses of (a) the difficulty ordering found within the data themselves, and (b) the structure of the items. Based on this work, four ordered lists were created from 32 of the 36 items (four items were deleted due to idiosyncratic response patterns). Thus, the critical index computed over these items is bounded by 0.0 and 4.99.

Synthetic Word Spelling. In this task, students were asked to spell synthetic words. The words were constructed to correspond to the orthography of the language assessed using the same criteria employed in constructing the materials for the Synthetic Word Decoding task. Each word was presented in isolation, and was read twice by the tester. Each student began the task with the first word, and continued to more difficult items unless he failed to respond to three items in a row.

For the 15 words presented, each of the items was scored in a manner similar to the scoring criteria used in the Synthetic Word Decoding task. A value of 3 was assigned to any item that was spelled without error. A value of 2 was given to those responses that were mostly correct (e.g., correct responses except for a minor letter-sound error such as a vowel shift within vowel family or a stress variant). A value of 1 was given for responses that were partly correct (e.g., correct responses except for a single vowel or consonant substitution or deletion). A value of 0 was assigned for assumed failure, no response, or for misspellings beyond those tolerated in the above categories. Note that the scoring was fairly stringent -- two major errors within an item were sufficient to receive a score of 0 (e.g., spelling sidded as sited). Note that the scoring was not based exclusively on the nominal letter-for-letter correspondence to the pronounced words, but on letter-sound correspondence (e.g., both glire and glier are correct spellings for the same synthetic word).

The index computed to represent performance on this task was the percentage of assigned points to total possible points (15 x 3 = 45). Again, this value cannot be tied to grade level expectations of performance.

For the English IRAS, one of the 15 words was replaced after the Year 2 administration (namely, vonning was replaced with feening) because of the difficulty of the vowel. No changes were made in the Spanish IRAS materials.

Sentence reading. In the next task, students were asked to read short two-sentence paragraphs as a way of both (a) assessing oral reading fluency, and (b) providing an efficient mechanism for placement in the comprehension tasks which followed. There were seven such paragraphs, each selected from the ordered narrative comprehension texts described below. Each student began with the first paragraph, and continued to more difficult ones if (a) the paragraph was read in 20 seconds or less (an average reading rate of 51 words per minute for the lowest ordered passage and a rate of 81 words per minute for the higher ordered passages), and (b) for three identified critical words, at least one was read correctly.

From this task, a summary index of oral reading fluency was obtained. In exploring the derivation of such a measure, the reading rates, in syllables per second, for each of the paragraphs successfully read was computed employing the English IRAS data from Year 4. At the level of the individual student, a best-fit regression line was computed through the data points available from the paragraph set read. The slopes of these lines were generally negative, indicating that reading rates declined as the difficulty of the material increased; further, the correlation between the predicted intercepts for Level A material and the actual data values at Level A was .97. Given this, the syllables per second measure computed on the easiest materials presented (Level A) was used as a summary index of oral reading fluency.

Based on analyses of the English Year 4 data, it was found that, on average, reading rates were constant for a given student over all paragraphs which were successfully read (i.e., in less than 20 seconds, and with two or fewer errors on the three critical words). Thus, a measure of reading fluency, syllables per second, for each student was computed by dividing the number of syllables contained on the Level A paragraph (18 in the English version and 29 in the Spanish version) by the time taken to read that paragraph.

This task was introduced in the Year 3 administration, and thus, given the cohort structure of the target sample, the reading fluency measure derived from it is not available in the early grades for students from Sites 0, 1, and 2.

For students who were not successful in Sentence Reading, the next task, Reading Comprehension, was not administered, again assuming failure since some skill in isolated sentence reading is necessary for success in reading connected text. For such students, the Listening Comprehension task was the next task administered.

Comprehension. In the next task, the student's reading comprehension was assessed. The materials consisted of nine levels, each containing two well-formed narrative and expository passages (except Level 1 which did not contain any exposition). Passages across levels were ordered in difficulty based on word frequency, number of words per sentence, number of sentences, and number of propositions expressed per sentence. Each story was constructed according to the principles of story grammar (Rumelhart, 1977), and associated with each element was a probe question.

The student entered this task at the level of highest success found in the sentence reading task described above. For the first four levels the student was presented with the appropriate narrative and asked to read it aloud. If the student was able to read the story in less than 150 seconds, then he was asked to retell as much of the story as he could. After the student finished the free recall task, any element that was not adequately recalled was then probed with the corresponding question. If the student met the reading time criterion, an expository text at the same level was presented, and comprehension was assessed using the free and cued recall procedures. This was followed by the presentation of more difficult levels until the time criterion for narrative reading was not met, or the highest level narrative for oral reading (level 4) had been given. If the student failed to meet the reading time criterion for the initial narrative presented, easier narrative and expository passages were presented until success was achieved.

For students who successfully met the reading time criterion at level 4, the next levels (5 to 9) were read silently. If half or more of the elements were successfully recalled in either the narrative or expository passages, at a given level, the next more difficult level was attempted.

For students who were not successful at Level 6 or higher, their listening comprehension was assessed for parallel narrative and expository passages read to them, again using the free and cued recall procedures. If the student recalled half or more of the passage elements under either free or cued recall, the next more difficult level was presented until the student failed to meet this criterion, or the highest level narrative and expository in the listening comprehension materials (Level 6) had been given. If the student failed to meet this criterion, listening comprehension of less difficult passages was assessed until the recall criterion was successfully met.

In scori comprehension components of the IRAS, each element under free anu cued recall was scored as "C", completely correct (all or most of the propositions expressed by the element were given correctly), "B", briefly mentioned (only some of the propositions expressed in the element were given correctly), "N", no response (none of the element's propositions were mentioned), or "W", incorrect response (the student's response was unrelated to the element's propositions). For any element receiving a "C" under free recall, its associated probe was not asked and was coded as an "S" (assumed success).

For passages not attempted because the recall criterion on a more difficult passage had been met, elements under free and cued recall were scored as "S", assumed success. For recall assessments not attempted because the recall criterion on a less difficult passage had not been met, elements were scored as "F", assumed failure. For students who failed to meet the reading time criterion for a given passage, recall was not assessed, and passage elements under such conditions were also coded as "F". The rationale for this procedure was that students reading at such slow rates would not be able to integrate sentence structures in a fast decaying short-term memory, and thus would fail to recall the elements adequately. If asked to do so, it was felt that the frustration from likely failure might impair performance on subsequent passages at a less difficult level.

After scoring elements separately for free and cued recall, each element was then assigned a single value ranging from 0 to 7, based on responses under both recall conditions as follows:

free recall value	cued recall value	combined value
S	S	7
C	S	7
B	C	6
B	B	4
B	N,W	3
N,W	C	5
N,W	B	2
N,W	N,W	0
F	F	0

Based on the element values above, critical indices were computed for both reading and listening comprehension, and for both narrative and expository passages under each of these conditions. For each, the integer portion of the index represented the level of highest success, based on meeting the criterion of recalling half or more of a passage's elements. The decimal portion of the value was the ratio of assigned points for combined elements to total possible points at the level of highest success.

Recall that students who successfully read passages at Level 7 (the highest listening comprehension level) or higher, were not assessed for listening comprehension. Under the assumption that their reading skill is not limited by their decoding skills, but only by their comprehension skills, the scores they received on reading comprehension were used as estimates of their listening comprehension skill.

In general, the English and Spanish IRAS comprehension texts covered the same content material, but changes were made in the Spanish versions where needed in order to maintain an appropriate vocabulary for the given level.

In the Year 1 administrations of both the English and Spanish versions of the IRAS, only three levels of texts were included. Within each level were three parallel narratives, one for assessing oral reading comprehension, one for silent reading comprehension, and one for listening comprehension. The administration procedure allowed for the assessment of oral reading comprehension, and, if successful, silent reading comprehension at the same level -- listening comprehension was assessed for each level beyond the level of highest success in oral reading comprehension. The recall procedures were slightly different from those discussed above, but still allowed individual story elements to be scored independently for free and cued recall responses. In deriving a critical index that was comparable across years, only the oral and listening comprehension segments were used, ignoring performance on silent reading comprehension. The critical indices for this year were thus bounded by 0.0 and 3.99, though few students reached these levels in either of the two comprehension assessments.

In the Year 2 administrations of both the English and Spanish versions of the IRAS, three additional levels were added to the three levels that comprised the Year 1 materials. There was a further adjustment to the recall procedures, but they still allowed independent assessments of free and cued recall for individual story elements. The critical indices computed in this year were bounded by 0.0 and 6.99, though few students reached these levels in either of the comprehension assessments.

In the Year 3 administrations, the separate silent reading texts were removed, and reading comprehension was assessed by having students read texts in a manner appropriate for the grade level of the material (i.e., Levels 1 to 4 were read aloud, and all higher levels were read silently). Expository texts were added to each level, except the lowest, for both reading and listening assessments. Further, an additional level of text was added for reading comprehension assessment (Level 7); assessments for listening comprehension were not made beyond Level 6. For the story set used in the previous two years, a modification in the probe questions associated with individual elements allowed some elements to be combined, thereby reducing the number of probes that needed to be asked (saving administration time), and concomitantly, reducing the number of elements scored; the stories themselves, however, remained unchanged.

The Year 4 administration was identical to the Year 3 administration. In the Year 5 testing, two additional text levels were included in the reading comprehension materials (Levels 8 and 9), and a small modification in the administration procedures was implemented in order to reduce testing time.

Metalinguistic tasks. During the third year of the IRAS administration, metalinguistic tasks were added to obtain information about how the students viewed their ability to perform certain tasks and to provide insights into how they might be accomplishing them. The metalinguistic tasks consisted of a few questions asked by the tester after successful completion of the tasks of Real Word Decoding ("How did you know that ____ was pronounced that way?"), Vocabulary Definitions ("What is a word? A sentence? A story?"), Sentence Reading ("If you were to break this sentence into parts, where would you make the breaks?"), and Reading Comprehension ("If the sentences in this passage were all jumbled, would the story make sense?"). The tape-recorded responses to these questions have not been fully analyzed, and will not be reported here.

Reliability

The reliability assessments of the IRAS summary measures for each scale by collection year are presented in Tables 19 through 23 for the

English version, and Tables 24 through 28 for the Spanish version. The analyses are based on the scale scoring procedures given above with two exceptions. First, given the scoring of Sentence Reading (i.e., syllables per second on the Level A paragraph), no reliability assessments were made. Second, the assessments for each of the four comprehension scales were made at the level of the component passages, rather than across the entire material set used in deriving the critical index summary measure of performance. Since the assessments described were made for each of the data collection years, the samples reflect the cohort structure described earlier (see Volume 2: Design of the Study).

For each of the ten tables which summarize the analyses, the scales are displayed down the lefthand column, with the relevant collection year assessments represented within each. For each yearly assessment, the number of cases, the number of items (i.e., the number of scale items less the number of items found to have no variance), the item total statistics (mean and standard deviation), the mean number of actual responses, and the reliability coefficient alpha (based on the number of actual responses) are presented. Concerning the coefficient of reliability, since each scale allowed assumed responses (for material that was either too easy or too difficult, based on the student's performance at testing time), the alpha coefficient was adjusted to reflect the total number of actual responses made. This adjustment procedure involves reducing the degrees of freedom associated with the residual error term by the number of assumed responses, recomputing the residual mean square based on the adjusted degrees of freedom, and then recomputing the alpha coefficient based on the adjusted residual mean square. In cases where there are an excessive number of assumed responses (an average actual response rate of one item or less per respondent), the procedure cannot be followed, and no coefficient is presented. Such only occurred in the individual passage assessments made within the comprehension tasks, and was the result of two general case types reflecting certain material-cohort combinations: (a) for higher-level materials presented to cohorts consisting of predominately early grade-level students (where much of the material resulted in assumed failure), and (b) for lower-level materials presented to cohorts consisting of predominately later grade-level students (where much of the material resulted in assumed success). In some cases, mostly in the Spanish IRAS, no reliability assessments were possible under any of the cohorts due to the performance levels of the respective students.

English administrations. Before discussing the results of the English IRAS reliability analyses, a review of the cohort structure will be presented. The description is idealized, in that it represents the sampling plan of the study and does not take attrition into account; the proportions of grade-level representation, however, should remain relatively stable.

For the English version, the students in the Year 1 sample represented the 20 first-grade bilingual students from the first cohort (all from Site 0). This group was supplemented by nine first-grade students

Table 19

Interactive Reading Assessment System - English:

Reliability Analysis of the Total Scale Scores for Non-Comprehension Scales for Each Collection Year

Scale	Collection Year	N of Cases	N of Items*	Item Total Mean	Item Total SD	Mean Number of Actual Responses	αk
Real Word Decoding	1	29	84	69.7	60.4	17.0	.92
	2	67	72	49.8	50.1	12.6	.88
	3	107	84	65.2	70.2	16.0	.94
	4	299	84	65.8	70.4	15.8	.94
	5	250	84	104.6	62.0	18.6	.93
Vocabulary Definition	1	29	48	50.4	32.5	15.0	.90
	2	68	54	48.0	35.0	19.1	.93
	3	105	42	35.7	26.7	11.3	.87
	4	299	42	39.0	28.0	12.0	.89
	5	249	42	54.8	20.5	10.7	.82
Synthetic Word Decoding	1	29	35	22.1	29.3	12.2	.94
	2	68	35	13.6	22.8	8.7	.89
	3	107	41	31.8	35.8	24.1	.97
	4	297	41	32.5	36.2	22.7	.97
	5	250	41	54.1	34.3	34.6	.96
Synthetic Word Spelling	1	29	12	7.5	6.9	3.2	.33
	2	67	15	4.5	6.2	5.2	.58
	3	104	15	8.6	9.7	11.1	.88
	4	297	15	8.3	9.7	9.2	.85
	5	244	15	13.0	9.6	13.6	.86

*Items with no variance were deleted from the analysis. For each deleted item, its mean was 0 (i.e., all respondents answered incorrectly).

Note: All scales allowed assumed success and failure, and the reliability coefficients were adjusted for the number of such "responses" by reducing the residual degrees of freedom proportionately, and then recomputing the residual mean square and coefficient alpha on which each was based.

Table 20

Interactive Reading Assessment System - English:

Reliability Analysis of the Total Scale Scores for Narrative Reading Comprehension Scales for Each Collection Year

Story Level	Collection Year	N of Cases	N of Items*	Item Total Mean	Item Total SD	Mean Number of Actual Responses	αk
A	1	29	7	21.8	16.9	5.8	.87
	2	68	7	15.0	19.1	2.2	.73
	3	107	4	12.6	13.1	0.7	-
	4	299	4	10.8	13.1	0.5	-
	5	251	4	21.5	11.4	0.4	-
B	1	29	11	17.2	27.4	3.8	.90
	2	68	11	20.8	30.1	2.8	.85
	3	107	6	15.9	19.0	1.2	.61
	4	299	6	15.2	19.0	1.0	-
	5	251	6	29.9	17.5	1.0	-
C	1	29	11	11.4	22.0	3.0	.77
	2	68	11	13.4	24.8	2.8	.81
	3	107	6	9.9	17.0	0.8	-
	4	299	6	11.0	17.4	0.8	-
	5	251	6	24.9	19.0	1.5	.85
D	2	68	13	5.1	18.7	0.8	-
	3	107	8	7.0	17.8	0.4	-
	4	299	8	8.4	18.4	0.8	-
	5	251	8	19.2	23.3	2.0	.87
E	2	68	10	1.0	8.0	0.1	-
	3	107	8	5.4	14.6	0.8	-
	4	299	8	6.5	15.7	1.0	-
	5	251	8	14.5	19.3	2.9	.86
F	2	68	7	0.5	4.1	0.1	-
	3	107	6	3.2	9.7	0.5	-
	4	299	6	2.9	9.4	0.5	-
	5	251	6	5.1	11.2	1.9	.75
G	3	107	8	2.0	6.9	0.7	-
	4	299	8	2.1	7.5	0.7	-
	5	251	8	3.3	8.6	1.7	.04
H	5	251	8	1.1	5.1	0.9	-
I	5	251	7	0.3	2.5	0.1	-

*Items with no variance were deleted from the analysis. For each deleted item, its mean was 0 (i.e., all respondents answered incorrectly).

Note: All scales allowed assumed success and failure, and the reliability coefficients were adjusted for the number of such "responses" by reducing the residual degrees of freedom proportionately, and then recomputing the residual mean square and coefficient alpha on which each was based.

Table 21

Interactive Reading Assessment System - English:

Reliability Analysis of the Total Scale Scores for Expository Reading Comprehension Scales for Each Collection Year

Story Level	Collection Year	N of Cases	N of Items*	Item Total Mean	Item Total SD	Mean Number of Actual Responses	αk
B	3	107	4	8.5	12.0	0.6	-
	4	299	4	8.7	12.0	0.6	-
	5	251	4	17.8	12.4	0.7	-
C	3	107	6	7.3	14.6	0.7	-
	4	299	6	8.9	16.1	0.7	-
	5	251	6	20.7	19.2	1.4	.73
D	3	107	6	4.5	12.4	0.2	-
	4	299	6	5.8	13.0	0.6	-
	5	251	6	12.6	16.8	1.4	.66
E	3	107	6	4.1	11.0	0.6	-
	4	299	6	4.1	10.8	0.8	-
	5	251	6	9.0	13.4	2.2	.74
F	3	107	6	2.8	9.4	0.4	-
	4	299	6	3.0	9.6	0.5	-
	5	251	6	5.7	11.2	1.7	.60
G	3	107	8	2.7	9.4	0.7	-
	4	299	8	2.9	10.2	0.7	-
	5	251	8	3.9	9.9	1.5	.01
H	5	251	8	1.5	5.9	0.9	-
I	5	251	6	0.2	1.9	0.1	-

*Items with no variance were deleted from the analysis. For each deleted item, its mean was 0 (i.e., all respondents answered incorrectly).

Note: All scales allowed assumed success and failure, and the reliability coefficients were adjusted for the number of such "responses" by reducing the residual degrees of freedom proportionately, and then recomputing the residual mean square and coefficient alpha on which each was based.

Table 22

Interactive Reading Assessment System - English:

Reliability Analysis of the Total Scale Scores for Narrative Listening Comprehension Scales for Each Data Collection Year

Story Level	Collection Year	N of Cases	N of Items	Item Total Mean	Item Total SD	Mean Number of Actual Responses	α_k
A	1	29	7	31.2	20.6	2.4	.84
	2	68	7	29.6	17.8	4.9	.87
	3	107	4	20.4	11.3	1.8	.78
	4	299	4	22.7	9.8	0.8	-
	5	251	4	26.7	5.8	0.1	-
B	1	29	11	28.0	37.0	0.4	-
	2	68	11	32.4	33.1	3.9	.91
	3	107	6	26.4	17.9	2.0	.86
	4	299	6	27.0	15.5	3.5	.87
	5	251	6	36.7	11.8	1.0	-
C	1	29	11	17.0	28.3	2.3	.75
	2	68	11	25.3	33.0	[illegible].8	.90
	3	107	6	21.0	18.2	2.0	.86
	4	299	6	25.3	16.9	2.5	.87
	5	251	6	35.7	13.4	1.0	-
D	2	68	13	19.1	30.0	3.6	.84
	3	107	8	23.8	23.5	3.1	.92
	4	299	8	28.1	22.9	3.8	.92
	5	251	8	41.0	21.4	2.4	.90
E	2	68	16	7.2	18.0	3.5	.56
	3	107	8	15.5	21.7	2.2	.89
	4	299	8	19.9	21.0	4.1	.91
	5	251	8	32.3	24.0	2.8	.92
F	2	68	7	2.1	6.9	0.7	-
	3	107	6	5.8	12.8	1.2	.34
	4	299	6	6.9	11.6	2.6	.81
	5	251	6	11.6	14.3	3.5	.89

Note: All scales allowed assumed success and failure, and the reliability coefficients were adjusted for the number of such "responses" by reducing the residual degrees of freedom proportionately, and then recomputing the residual mean square and coefficient alpha on which each was based.

Table 23

Interactive Reading Assessment System - English:

Reliability Analysis of the Total Scale Scores for Expository Listening Comprehension Scales for Each Collection Year

Story Level	Collection Year	N of Cases	N of Items	Item Total Mean	Item Total SD	Mean Number of Actual Responses	αk
B	3	107	4	15.1	12.5	1.2	.57
	4	299	4	15.2	11.3	2.3	.86
	5	251	4	22.4	10.1	0.7	-
C	3	107	6	18.1	18.2	1.9	.86
	4	299	6	18.4	17.3	2.5	.88
	5	251	6	29.9	17.7	1.0	-
D	3	107	6	9.5	14.0	2.4	.82
	4	299	6	12.5	16.0	2.8	.91
	5	251	6	23.8	19.1	1.8	.91
E	3	107	6	5.5	13.2	1.7	.91
	4	299	5	6.8	13.9	3.0	.95
	5	251	6	19.0	20.0	1.9	.95
F	3	107	6	5.8	12.7	1.2	.06
	4	299	6	7.0	11.7	2.6	.77
	5	251	6	13.7	15.0	3.3	.87

Note: All scales allowed assumed success and failure, and the reliability coefficients were adjusted for the number of such "responses" by reducing the residual degrees of freedom proportionately, and then recomputing the residual mean square and coefficient alpha on which each was based.

341

Table 24

Interactive Reading Assessment System - Spanish:

Reliability Analysis of the Total Scale Scores for Non-Comprehension Scales for Each Collection Year

Scale	Collection Year	N of Cases	N of Items*	Item Total Mean	Item Total SD	Mean Number of Actual Responses	α_k
Real Word Decoding	1	29	84	46.8	74.6	11.6	[illegible]
	2	67	84	33.3	63.4	11.0	.93
	3	159	84	85.2	92.4	22.7	.98
	4	300	84	81.4	100.0	16.1	.98
	5	257	84	98.1	99.5	20.6	.98
Vocabulary Definition	1	29	50	48.0	39.5	18.2	.95
	2	49	56	52.1	39.0	33.7	.95
	3	132	42	41.9	34.7	11.9	.92
	4	298	42	38.8	38.0	11.0	.94
	5	243	42	47.8	38.0	11.6	.94
Synthetic Word Decoding	1	29	32	22.2	32.6	10.9	.96
	2	65	32	14.7	27.7	7.8	.94
	3	159	32	36.0	35.1	18.8	.97
	4	295	32	31.9	35.7	17.9	.98
	5	257	32	39.4	34.9	21.8	.98
Synthetic Word Spelling	1	29	15	9.7	13.2	6.2	.91
	2	67	15	6.8	11.0	4.9	.85
	3	157	15	12.3	11.7	11.1	.90
	4	298	15	11.0	12.7	8.8	.91
	5	252	15	10.5	13.3	10.5	.92

*Items with no variance were deleted from the analysis. For each deleted item, its mean was 0 (i.e., all respondents answered incorrectly).

Note: All scales allowed assumed success and failure, and the reliability coefficients were adjusted for the number of such "responses" by reducing the residual degrees of freedom proportionately, and then recomputing the residual mean square and coefficient alpha on which each was based.

Table 25

Interactive Reading Assessment System - Spanish:

Reliability Analysis of the Total Scale Scores for Narrative Reading Comprehension Scales for Each Collection Year

Story Level	Collection Year	N of Cases	N of Items*	Item Total Mean	Item Total SD	Mean Number of Actual Responses	α_k
A	1	29	7	11.4	17.5	2.2	.77
	2	67	7	6.7	15.4	0.9	-
	3	160	4	5.4	9.5	0.8	-
	4	300	4	6.9	11.7	0.3	-
	5	257	4	10.7	13.3	0.3	-
B	1	29	11	7.0	18.2	1.5	-
	2	67	11	9.8	21.9	1.8	.50
	3	160	6	5.0	12.4	0.7	-
	4	300	6	8.5	16.0	0.7	-
	5	257	6	14.8	19.1	0.9	-
C	1	29	10	2.3	12.6	0.3	-
	2	67	11	4.4	15.7	0.8	-
	3	160	6	1.8	7.7	0.3	-
	4	300	6	5.9	14.2	0.4	-
	5	257	6	11.8	18.2	0.6	-
D	2	67	10	0.9	7.0	0.1	-
	3	160	8	0.8	6.4	0.1	-
	4	300	8	5.3	15.4	0.5	-
	5	257	8	10.2	20.5	0.7	-
E	2	67	0	0.0	0.0	0.0	-
	3	160	8	0.7	6.2	0.0	-
	4	300	8	4.3	13.0	0.8	-
	5	257	8	8.0	17.0	1.4	.56
F	2	67	0	0.0	0.0	0.0	-
	3	160	[illegible]	0.3	3.4	0.0	-
	4	300	6	1.1	5.0	0.4	-
	5	257	6	2.8	8.7	0.8	-
G	3	160	8	0.3	2.5	0.1	-
	4	300	8	1.2	5.3	0.4	-
	5	257	8	1.7	6.3	0.8	-
H	5	257	8	0.4	2.4	0.4	-
I	5	257	1	0.0	0.0	0.0	-

*Items with no variance were deleted from the analysis. For each deleted item, its mean was 0 (i.e., all respondents answered incorrectly).

Note: All scales allowed assumed success and failure, and the reliability coefficients were adjusted for the number of such "responses" by reducing the residual degrees of freedom proportionately, and then recomputing the residual mean square and coefficient alpha on which each was based.

343

Table 26

Interactive Reading Assessment System - Spanish:

Reliability Analysis of the Total Scale Scores for Expository Reading Comprehension Scales for Each Collection Year

Story Level	Collection Year	N of Cases	N of Items*	Item Total Mean	Item Total SD	Mean Number of Actual Responses	αk
B	3	160	4	2.8	7.8	0.4	-
	4	300	4	5.1	10.0	0.4	-
	5	257	4	8.7	12.2	0.5	-
C	3	160	6	1.0	5.8	0.2	-
	4	300	6	4.9	12.7	0.4	-
	5	257	6	8.9	16.7	0.3	-
D	3	160	6	0.5	4.7	0.0	-
	4	300	6	3.5	10.3	0.5	-
	5	257	6	6.5	14.1	0.5	-
E	3	160	6	0.5	4.7	0.0	-
	4	300	6	2.5	8.5	0.5	-
	5	257	6	4.8	10.9	1.1	-
F	3	160	6	0.4	4.0	0.0	-
	4	300	6	1.5	6.2	0.4	-
	5	257	6	2.9	9.0	0.8	-
G	3	160	7	0.4	3.3	0.1	-
	4	300	8	1.6	7.3	0.4	-
	5	257	8	2.1	7.6	0.8	-
H	5	257	7	0.5	2.6	0.3	-
I	5	257	2	0.0	0.4	0.0	-

*Items with no variance were deleted from the analysis. For each deleted item, its mean was 0 (i.e., all respondents answered incorrectly).

Note: All scales allowed assumed success and failure, and the reliability coefficients were adjusted for the number of such "responses" by reducing the residual degrees of freedom proportionately, and then recomputing the residual mean square and coefficient alpha on which each was based.

Table 27

Interactive Reading Assessment System - Spanish:

Reliability Analysis of the Total Scale Scores for Narrative Listening Comprehension Scales for Each Data Collection Year

Story Level	Collection Year	N of Cases	N of Items*	Item Total Mean	Item Total SD	Mean Number of Actual Responses	α_k
A	1	[illegible]	7	36.2	11.5	6.0	.68
	2	67	7	33.9	12.1	5.7	.68
	3	160	4	21.7	7.8	3.0	.74
	4	300	4	17.7	12.2	1.6	.85
	5	257	4	20.8	11.0	1.4	.68
B	1	29	11	11.9	24.3	1.9	.56
	2	67	11	29.3	29.9	6.1	.92
	3	160	6	24.1	14.1	4.4	.87
	4	300	6	18.9	17.5	3.7	.94
	5	257	6	23.4	18.0	2.3	.87
C	1	29	10	3.4	13.4	0.7	-
	2	67	11	26.9	30.1	4.6	.90
	3	160	7	20.7	19.4	3.6	.91
	4	300	7	19.6	20.7	2.6	.90
	5	257	7	26.6	21.8	2.0	.88
D	2	67	13	17.3	24.0	5.0	.80
	3	160	8	21.1	21.2	4.3	.91
	4	300	8	18.0	22.6	2.8	.90
	5	257	8	25.7	24.2	3.0	.91
E	2	67	11	4.5	12.7	1.6	-
	3	160	8	7.6	15.1	1.8	.65
	4	300	8	13.3	19.8	2.4	.85
	5	257	8	19.0	23.2	2.7	.92
F	2	67	6	1.4	5.6	0.4	-
	3	160	6	1.9	6.3	0.8	-
	4	300	6	4.5	10.1	1.5	.32
	5	257	6	6.5	11.9	2.1	.72

*Items with no variance were deleted from the analysis. For each deleted item, its mean was 0 (i.e., all respondents answered incorrectly).

Note: All scales allowed assumed success and failure, and the reliability coefficients were adjusted for the number of such "responses" by reducing the residual degrees of freedom proportionately, and then recomputing the residual mean square and coefficient alpha on which each was based.

Table 28

Interactive Reading Assessment System - Spanish:

Reliability Analysis of the Total Scale Scores for Expository Listening Comprehension Scales for Each Collection Year

Story Level	Collection Year	N of Cases	N of Items	Item Total Mean	Item Total SD	Mean Number of Actual Responses	αk
B	3	160	4	14.7	9.7	2.8	.79
	4	300	4	11.2	14.5	2.5	.89
	5	257	4	14.4	12.3	1.5	.78
C	3	160	6	15.1	15.3	3.0	.86
	4	300	6	13.2	16.3	2.2	.87
	5	257	6	18.1	18.6	1.7	.84
D	3	160	6	7.4	9.8	3.0	.73
	4	300	6	7.4	13.6	2.0	.85
	5	257	6	12.9	17.1	2.2	.91
E	3	160	6	1.5	6.1	1.1	-
	4	300	6	4.6	12.1	1.8	.90
	5	257	6	7.6	15.2	2.0	.94
F	3	160	6	2.1	7.0	0.6	-
	4	300	6	4.5	10.5	1.4	.47
	5	257	6	6.7	11.9	2.1	.71

Note: All scales allowed assumed success and failure, and the reliability coefficients were adjusted for the number of such "responses" by reducing the residual degrees of freedom proportionately, and then recomputing the residual mean square and coefficient alpha on which each was based.

from the same classrooms in order to gain a larger sample for assessing the then newly-developed IRAS instrument package. In Year 2, the sample consisted of 50 first-grade students and 20 second-grade students, all from the border sites. The Year 3 sample was also exclusively composed of the border site students, with 50 first-graders, 50 second-graders, and 20 third-graders. In Year 4, the sample contained the same border site students as in Year 3 (at their next grade level), but also contained, from the non-border sites, 160 bilingual and 40 monolingual-English first-grade students. In the final year, the sample matched that of Year 4, following students into their next instructional year, with the exception of the 20 Year 4 fourth-grade students from Site 0 who exited the study.

Given this structure, no two samples across the collection years are strictly comparable -- the closest two are those of the last two years, but as seen in the actual sample sizes, the attrition rate reveals a rather substantial influence, leaving their comparability in doubt. Thus, while sample size within a collection year is relatively stable across scales, comparisons of the descriptive statistics between years are difficult to interpret, and will not be made in the following discussions.

The results of the assessments are presented in Tables 19 through 23. Table 19 summarizes the four non-comprehension scales of Real Word Decoding, Vocabulary Definition, Synthetic Word Decoding, and Synthetic Word Spelling. Tables 20 through 23 summarize the four comprehension scales, Narrative Reading, Expository Reading, Narrative Listening, and Expository Listening, respectively.

For the non-comprehension scales presented in Table 19, it is important to remember the changes in materials over the collection years, most notably: (a) in Year 3, the reduction from four to three words per list in Vocabulary Definition (changing the total number of items from 56 to 42), and (b) the introduction of a new set of synthetic words in the Year 3 administration of Synthetic Word Decoding (changing the total number of items from 36 to 42).

The average total scale scores across tasks and collection years translate into average performance values ranging from 10% to 44%, giving some evidence, in the aggregate, of a floor effect for some scales (most notably, in the synthetic word tasks), but little evidence of ceiling effects. For Real Word Decoding and Vocabulary Definition, material sets that are directly comparable, performance in the latter, as expected, exceeds that in the former. The reliability coefficients for the first three tasks are all quite high, ranging from .82 to .97. For the first two years of assessments of the fourth task, Synthetic Word Spelling, the coefficients are low, due to the low number of actual responses, but for the next three years, they range from .85 to .88 as the number of actual responses increases.

The reliability assessments for the nine passages appearing in Narrative Reading Comprehension over the five years of data collection are summarized in Table 20. First, note that many of the coefficients are missing, for reasons discussed earlier; recalling the cohort structure, the pattern of missing coefficients can be summarized as follows. First, for the lower-level stories, actual response rates are highest from the first collection years since the cohorts tested then show the largest proportions of early-grade students for whom such materials are appropriate -- thus, assessments are possible for the initial cohorts, but not for later cohorts where the sample proportions do not largely favor the younger students. The mid-level stories (a) are too difficult for the early-grade students (thus, the actual response rate is low in the first years of administration), (b) are appropriate for the middle-grade students (the actual response rate is higher for the middle years of administration), and (c) are too easy for the later-grade students (the actual response rate becomes low again in the later administration years). Accordingly, the assumed response rate is highest in the early and later year administrations, allowing reliability assessments only in the mid-year administrations. For the most difficult material, assessments are only possible in the last years of administration (if at all) where the largest proportion of later-grade level students appears. This general pattern holds for all comprehension assessments in both English and Spanish.

Second, note that within each administration year, performance declines with increases in material difficulty, as would be expected given the structure of the material set. Further, within each story level, the alpha coefficient is largest where the number of actual responses is highest -- for Levels A through E, these coefficients are quite acceptable (the largest ones within each level ranging from .85 to .90), while the reliability of the most difficult material, Levels F through I, could not be adequately assessed with the sample of students available.

Table 21 displays the results of the assessments of the Expository Reading Comprehension materials. Recall that these materials were first introduced in the Year 3 administrations, and supplemented in Year 5, thus lessening the opportunities to assess their reliability. The general description of the assessments made, however, follows that given for Narrative Reading Comprehension. First, within administration years, performance declines with increases in material levels. Second, where the actual response rates are sufficiently high, the alpha coefficients are acceptable, though neither the response rates nor the coefficients are as large as those found in Narrative Reading Comprehension -- the samples tested do not allow sufficient assessments of the materials at the extremes (Level B; and Levels G, H, and I).

Tables 22 and 23 summarize the assessments made on the material sets for Narrative and Expository Listening Comprehension, respectively. Again, performance within collection years declines with increases in material difficulty. Second, the alpha coefficients increase with increases in the actual response rate, and within each level, the largest coefficients are quite high (ranging from .87 to .92 for the narrative materials, and from .86 to .95 for the expository materials). Given that performance on listening comprehension generally exceeds that of reading comprehension (especially in the elementary grades), reliability assessments were possible on all of the listening comprehension passages. Further, although the demands of the tasks are quite different, the strong reliability of the listening comprehension materials suggest the reliability of the reading comprehension materials, had the sample allowed such an assessment, since the passages for each were constructed in a parallel fashion.

In summary, where the material sets could be assessed, the coefficients obtained strongly support the reliability of the summary measures derived. Next, consideration will be given to the Spanish version reliabilities.

Spanish administrations. The cohort structure represented in the Spanish administrations matched that of the English administrations for the set of bilingual students (see above). None of the monolingual-English students were assessed with the Spanish version, however, the 40 monolingual-Spanish students (who were not assessed with the English version) were included in the Spanish administrations for Years 3 through 5, as they passed from first through third grade, respectively.

Tables 24 through 28 present the results of the reliability analyses for the Spanish tasks. In Table 24, the assessments of the non-comprehension scales of Real Word Decoding, Vocabulary Definition, Synthetic Word Decoding, and Synthetic Word Spelling are summarized. First, recall that the English and Spanish samples assessed during the first two years are identical, thus permitting performance to be compared across the two language versions (save some differences in missing data within scales which are relatively inconsequential except in Year 2 Vocabulary Definition) -- comparisons cannot be made for the last three years due to the differential inclusions of the monolingual samples. In such comparisons, average performance is very similar, as expected given the bilinguality of the sample. Across all scales and years, the range of average performance resembles that found in the corresponding English scales (from about 10% to 45%), again showing no signs of a ceiling effect in the aggregate, but perhaps some evidence of floor effects (again, especially in the synthetic word tasks). Over all administration years, the number of actual responses is high enough to allow sufficient assessments of reliability. As seen previously in the English scales, the coefficients for a given scale increase with increases in the number of actual responses; all obtained coefficients are high (the lowest is .85), supporting the reliability of the summary measures derived.

Tables 25 through 28 display the results of the comprehension analyses (Narrative Reading, Expository Reading, Narrative Listening, and Expository Listening, respectively). For Narrative Reading (Table 25), only the Level A story could be sufficiently assessed (Year 1, with an alpha coefficient of .77), as the number of actual responses is too low in all other cases. This is reflective of each sample's relatively poor performance, in the aggregate, in Spanish reading. For Expository Reading (Table 25), the response rates would not allow any of the material levels to be assessed.

For Narrative and Expository Listening (Tables 27 and 28, respectively), aggregate performance is clearly superior to that of reading, at all material levels, and for all assessment years. The actual response rates are sufficiently high to allow each of the levels to be assessed; the highest coefficients within each level (generally corresponding to the highest actual response rates) range from .72 to .92 for the narratives, from .71 to .94 for exposition. Again, the high reliabilities found for the listening material sets, coupled with the parallel construction of the listening and reading passages, supports the contention that the reading material sets would have been found reliable had the sample possessed sufficient skill to allow such assessments.

In summary, the coefficients obtained strongly support the reliability of the scales and their derived summary measures. In those cases where coefficients could not be obtained, mostly in the reading comprehension tasks, and more frequently in the Spanish version than in English, the performance of the sample simply was not sufficient to allow such assessments.

Descriptive Statistics

In this section, the longitudinal performance of the bilingual sample will be described, first over the English scales, then the Spanish scales. The focus will not be on the actual measures derived year by year, but on the linear growth functions computed from such yearly data. These functions, completely captured by the slope and intercept (either Y or X), were discussed earlier (see Volume 3: Measurement of Growth), but will be briefly reviewed here. First, the rationale for their use will be given, followed by a description and interpretation of the actual measures computed, and a discussion of an adjustment procedure employed to deal with a certain type of floor-ceiling problem. Finally, the descriptive data on the growth measures for the bilingual sample will be presented.

The rationale behind computing the linear indices of growth over the IRAS scales is three-fold. First, many of the IRAS material sets were designed around a linear progression in readability, and thus, a linear component is inherent in each of the so-designed tasks. Second, as was demonstrated for the aggregate measures discussed in Volume 3, much of the growth of the study's sample can be "explained" as linear -- to be sure, there are interesting departures from such linearities for individual students, and the degree to which these can be accounted for by the associated individual instructional data is of great import, and will be treated in a subsequent volume. Nonetheless, the strong linear trends in the data themselves argue that a substantial degree of variability can be accounted for by a linear model of growth. Third, given that the data generally reveal strong linear trends irrespective of the number of data points on which the trends are based, a straightforward solution to the difficult problem of missing data (faced by any longitudinal study) is at hand. For these reasons, in this study, growth in reading is mainly depicted as linear growth (with associated indices of nonlinearity also computed) -- the derivation of these measures is described below.

Linear estimates. For each individual student, the slope and intercept of the best-fitting line were computed for each of the nine IRAS scales in both English and Spanish. The slope is easily interpretable as the best estimate of (linear) growth in the relevant skill area resulting from a single year of instruction. However, as stated in Volume 3, there are substantive issues regarding the point from which growth should be measured. As argued there, the issue is not mathematical, as any arbitrary point can be taken along the X-axis. Rather, the question is what point is most appropriate given those experiences of the student (whether at home or in school) that would begin to advance literacy from some zero value. For this study, the intercept at first grade was used, since this point would seem to be the modal value of the sample's first systematic instruction in literacy -- this is clearly an estimate, but the study does not possess the pre-kindergarten data that would be necessary to accurately determine this placement.

An alternative intercept was also computed -- the X-intercept (which is the ratio of intercept to slope, multiplied by -1). This intercept (the point at which the best-fitting line crosses the X-axis) represents the estimated point at which the student would first show some skill on the IRAS task for which the data are being fit. Assuming that (a) the development of a component reading skill begins only when some appropriate effective instruction begins, and (b) the relevant IRAS task is then capable of detecting the resultant development in skill, this intercept can be interpreted as an estimate of the onset of effective instruction in the component skill area. Such estimates can range from relatively large positive values (representing a delay in school-provided effective instruction) to relatively large negative values (representing early effective instruction which is not school-based). For these estimates, values smaller than -5.0 (i.e., estimates that instruction began more than five years prior to kindergarten entry) are difficult to interpret. Such outliers were reset to -5.0 in order to reduce their influence on any means computed over the estimates.

Note that the interpretations of the Y- and X-intercepts are related: for any line with a positive slope, if the Y-intercept (at X = 0) is positive, then the X-intercept will be negative -- the interpretation of the Y-intercept is that the student began school with some (estimable) advantage; for the X-intercept, the interpretation is that the student encountered some instruction in the component skill at some (estimable) point prior to schooling. If the Y-intercept (again, at X = 0) is negative for a line with positive slope, then the X-intercept will be positive -- these allow the interpretations of (a) beginning schooling in the component skill area with some disadvantage (the magnitude estimated by the Y-intercept) or (b) delaying effective instruction in the component skill area until some point after school entry (estimated by the X-intercept). Given these interpretations, the

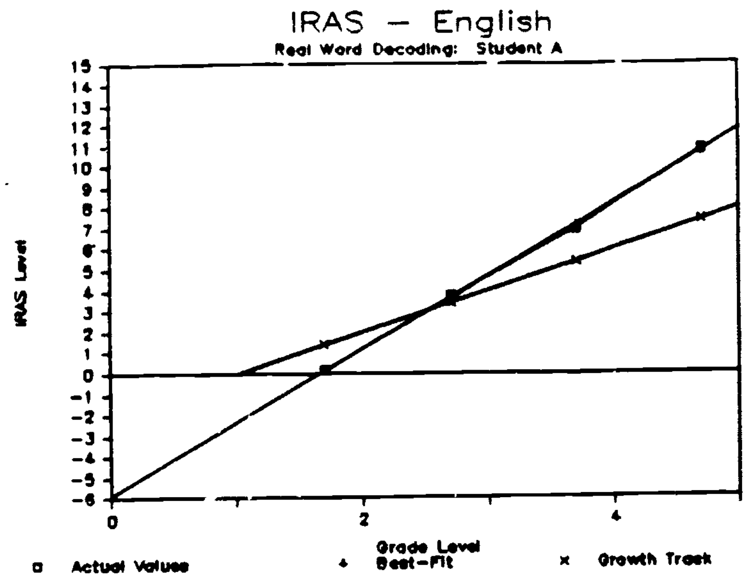
Y- and X-intercepts will be referred to as the student-intercept and the instruction-intercept, respectively.

As an index of nonlinearity, a standard measure of the amount of variability around the best-fit line was computed (and then converted to a percentage): one minus the r-squared value (the latter, of course, expressing the squared correlation coefficient between scale performance and grade level). Its interpretation is straightforward. At the extreme of 100%, the data show no linear component (i.e., the average of the data values is the best estimate of performance for any arbitrary grade-level value). At the other extreme of 0% unexplained variance, each of the data values falls precisely on the projected line, and scale performance is perfectly predictable for any arbitrary grade-level value. For values in between the two extremes, some linear component is present in the data -- as the percentage of unexplained variance decreases, so does the average (linear) prediction error.

Before providing some examples employing these measures, a few special cases should be noted with regard to the linear parameters computed. First, for students whose individual critical indices show no change over time, the slope of the best-fit line is zero (as it parallels the X-axis); the Y-intercept (at any X-value) is equal to the mean of the data points (which, of course, is equivalent to any of the individual points); the X-value is undefined formally, but was set to 0; and the unexplained variance is 100%. For students with two data points, the slope and intercept values present no special problems, but the unexplained variance is necessarily 0% (unless the two data points are equal, in which case the unexplained variance is 100%). For the discussions of the unexplained variance which appear in subsequent sections, these special cases of no-growth and two data points will be particularly important.

To illustrate the parameters discussed above, consider the two graphs presented in Figure 11 (a reproduction of Figure 3 from Volume 3). The two panels display the actual critical indices for three students on Real Word Decoding (identified as Vocabulary Decoding in Volume 3) over the four years in which they were tested, along with the best-fit line projected through these points, and the linear growth track. In the top panel, Student A reveals almost perfect linear growth -- the unexplained variance around the best-fit line is 0.08%. The rate of growth, 3.51 IRAS levels per instructional year, exceeds that expected from the growth track model (2 IRAS levels per year). The first-grade intercept is at -2.31, and suggests that the student began first-grade at a relative disadvantage; an alternative interpretation is provided by the X-intercept, which has a value of 1.66, and suggests that effective instruction did not commence until about two-thirds of the way through first grade.

The data from two additional students are displayed in the bottom panel of Figure 11. The data of Student B evidence a large degree of nonlinearity, which is reflected in the value of 17.90% unexplained variance. Growth approximates that expected by the growth track with a slope of 2.21. The first-grade Y-intercept of 4.42 suggests that the



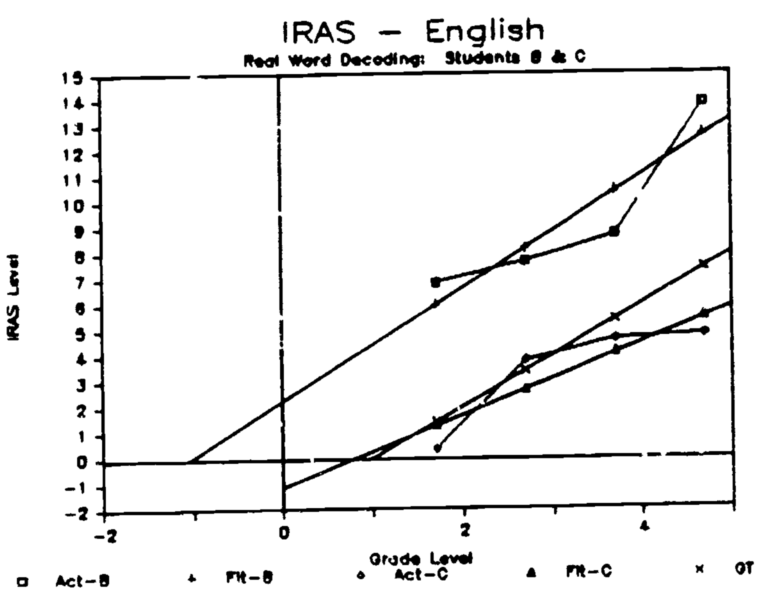


Figure 11. IRAS-E Real Word Decoding actual values and best-fit regression lines for three students.

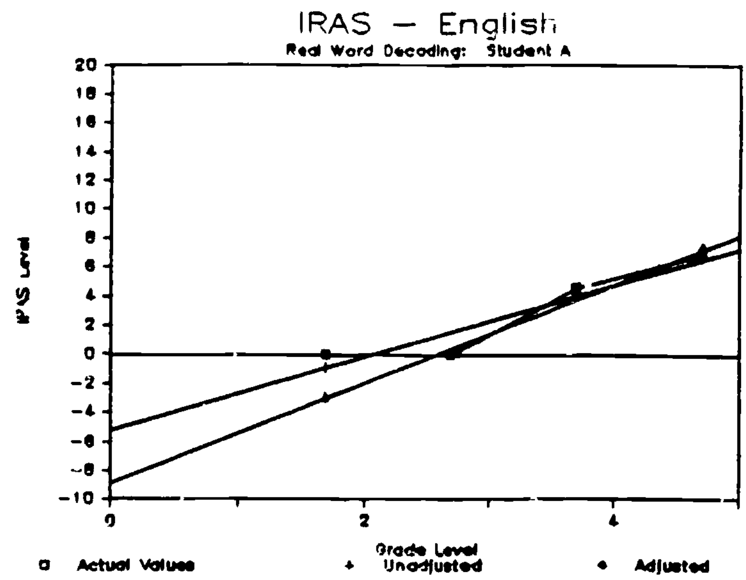
student may have had significant exposure to print prior to first grade (either in the home, during kindergarten, or both) or may have experienced a particularly effective first-grade instructional program. Alternatively, the X-intercept value of -1.00 suggests that such effective literacy instruction began a year before the student entered kindergarten.

The data from Student C, also displayed in the bottom panel of Figure 11, again, show a definite degree of nonlinearity, with 23.08% of the variance unexplained by the linear fit. This student shows less growth than that expected from the growth track model (the slope of the best-fit line is 1.41), but the first-grade Y-intercept value of 0.28 is in line with that predicted by the model. Note that the X-intercept, with a value of 0.80, suggests that effective instruction in decoding real words in English began near the beginning of first grade, again, as expected from the growth track model.

Floor-ceiling adjustments. As suggested by the descriptive data associated with the reliability analyses discussed above, there is some evidence, in the aggregate, of floor effects for some of the measures. However, both floor and ceiling effects can be found when examining patterns of growth for individual students. If, for a given student, the testing materials are not sensitive enough to detect any skill on a given task at initial testing, but do detect growth over the ensuing years, the question arises: What is the best estimate of growth, given the early insensitivity of the instrument? If a best-fit line is projected through the data points available, then it can be argued that the slope of such a line underestimates the actual growth of the student. Similarly, if a given student is successful at the highest levels of a given task, and again, a best-fit line is projected through the data points available, can this be said to be an accurate portrayal of the student's growth?

In this study, it was felt that such estimates of growth were sufficiently problematic to seek some solution. As a result, a procedure was followed whereby whenever successive initial floor or final ceiling effects were found for an individual student, these points would be ignored in computing the growth functions. Note, that the solution is not completely satisfactory, as a single initial floor or ceiling data point encountered over the sequence of testing cannot be adjusted -- for such cases, the test simply cannot provide the data needed to yield a more accurate estimate of skill. A few cases will illustrate the procedure and its effects.

Figure 12 displays the performance of two students (different from those depicted in Figure 11) on the English IRAS Real Word Decoding task. Recall that this task contains 14 levels, and is therefore bounded by 0.0 (no success on any of the lists) and 14.99 (complete mastery of the highest level). The top panel of the figure displays an instance of a floor effect on this task for one of the students in the bilingual sample. The 'box-symbol' line of the figure traces the actual critical indices computed for the student over the four years of testing. The '+' line marks the best-fit line when considering all



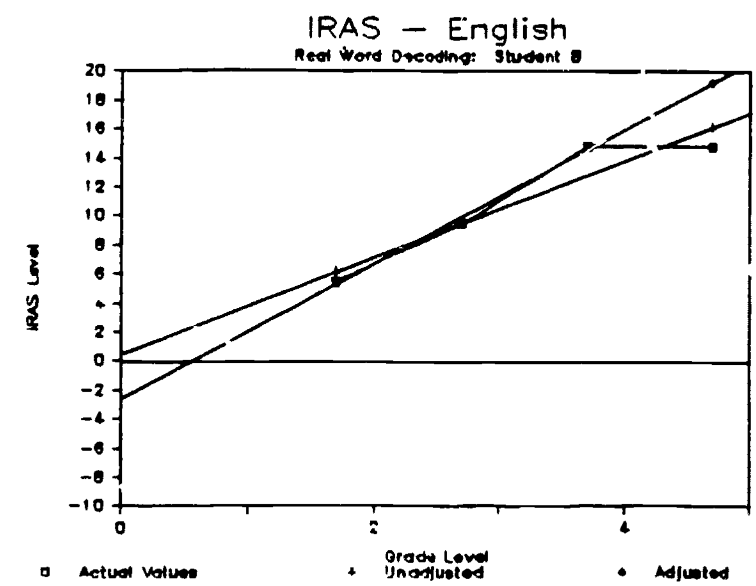


Figure 12. IRAS-E Real Word Decoding actual values and best-fit regression lines (both unadjusted and adjusted) for two students.

four data points -- it has a slope of 2.52 and a Y-intercept (at first grade) of -2.69. The 'diamond-symbol' line represents the best-fit line projected through the remaining three data points once the initial floor point has been ignored -- it has a slope of 3.45 and a Y-intercept of -5.48.

In the second panel, the data for a bilingual student showing a ceiling effect on the same task is presented. Again, the 'box-symbol' line follows the actual critical indices over the four-year testing sequence. The '+' line is the best-fit line when all four data points are considered in the computation -- it has a slope of 3.33 and a first-grade Y-intercept of 3.83. The 'diamond-symbol' line represents the best-fit line when considering only the initial three data values, ignoring the final point at the test ceiling -- it has a slope of 4.66 and a first-grade Y-intercept of 2.01.

As can be seen in both cases, the effect of the adjustment procedure is to increase the estimate of the slope and, concomitantly, lower the estimate of the Y-intercept (thereby increasing the X-intercept estimate) -- these are general effects for adjustments to positively sloped lines. Further, it clearly reduces the number of data points available for estimating growth. However, where the number of data points directly concerned the reliability of the estimates obtained from the best-fit line when considering differences between cohorts (see Volume 3), such is not the case here -- the reduction of data points under this procedure concerns only values which, for all practical purposes, are uninformative, and, we feel, detract from the reliability of the obtained estimates. There is, however, a separate issue regarding estimate reliability, namely, whether or not growth would continue to be linear had additional materials been added to the relevant tasks. This question simply cannot be answered in this study.

On the basis of these arguments, the adjustment procedure was employed for each of the IRAS measures in both English and Spanish. The floor was taken to be complete failure on the easiest level of material for each scale. The ceiling cutoffs, representing success on the highest level of the materials presented, are tabled below:

Scale	Ceiling Cutoff
Real Word Decoding	14.5
Vocabulary Definition	14.5
Synthetic Word Decoding	6.5 *
Synthetic Word Spelling	93.0
Sentence Reading	----
Narrative Reading	9.5
Expository Reading	9.5
Narrative Listening	9.5
Expository Listening	9.5

* 4.5 for Spanish

The number of cases where such adjustments were made was quite small (with the exception of the Spanish Reading Comprehension scales, where many students showed floor effects), but we believe that the growth measures derived are much more reliable estimates than those that would have been obtained had the adjustments not been made.

The discussion which follows treats the descriptive data for the growth measures computed over the English and Spanish IRAS scales for the bilingual sample.

English overall performance. In Tables 29 and 30, the overall performance of the bilingual sample is displayed for the non-comprehension and comprehension scales, respectively. The left margin is defined by the nine scales (five non-comprehension and four comprehension), and for each, the measures of slope, Y-intercept at first-grade (or student-intercept -- S-Intrcp), and X-intercept (or instruction-intercept -- I-Intrcp) are listed. For each of these, the descriptive measures of mean (M) and standard deviation (S) are provided. The number of cases (N) is given only for the slope measures, as it is identical for the two associated intercept measures listed. The task names are mnemonic, and stand for the following scale names:

VDC: Real Word Decoding (or Vocabulary Decoding)
VDF: Vocabulary Definitions
LDC: Synthetic Word Decoding (or Letter-sound Decoding)
LSP: Synthetic Word Spelling (or Letter-sound Spelling)
SRD: Sentence Reading
NRC: Narrative Reading Comprehension
ERC: Expository Reading Comprehension
NLC: Narrative Listening Comprehension
ELC: Expository Listening Comprehension

The single letter appended to each of the task names gives the language of the test, as English or Spanish (E or S, respectively).

Along the top of the table, the first column provides the descriptive statistics over the entire bilingual sample (overall), and then successively for students in the language entry categories of low English (Low Eng), high English (High Eng), low Spanish (Low Span), and high Spanish (High Span). These are then followed by a further refinement of language category based on combined English and Spanish skill: low English and low Spanish (Lo Lo), low English and high Spanish (Lo Hi), high English and low Spanish (Hi Lo), and high English and high Spanish (Hi Hi). These entry categories have been described elsewhere and will not be reviewed here (see Volume 4: Oral Language Growth for a detailed discussion, or, for a brief review in this volume, the English administration section under Assessment of Pre-reading Skills).

The data for individual sites (Sites 0, 1 2, 3, and 5) have been tabled and appear in Appendix B, with the first table of a pair providing the data for the non-comprehension scales, the second, the data for the comprehension scales. For these tables, the overall site data are provided first (Overall), followed by the data for the four

Table 29

Interactive Reading Assessment System - English:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Non-comprehension Scales Overall and by Language Category

Scale	Measure	Statistic	Overall	Low Eng	High Eng	Low Span	High Span	Lo Lo	Lo Hi	Hi Lo	Hi Hi
VDCE	Slope	M	3.2	2.8	3.5	3.2	3.2	2.5	3.1	3.9	3.3
VDCE	Slope	S	2.1	1.9	2.1	2.2	2.0	1.8	2.0	2.3	2.0
VDCE	Slope	N	246	113	133	111	135	59	54	52	81
VDCE	S-Intrcp	M	-0.1	-0.9	0.6	-0.7	0.4	-0.9	-0.9	-0.4	1.2
VDCE	S-Intrcp	S	2.9	2.5	3.0	2.7	3.0	2.3	2.7	3.0	2.9
VDCE	I-Intrcp	M	0.7	1.1	0.3	0.8	0.7	1.0	1.2	0.7	0.3
VDCE	I-Intrcp	S	1.4	1.2	1.5	1.6	1.3	1.5	0.9	1.7	1.4
VDFE	Slope	M	1.8	2.2	1.5	1.6	2.0	1.8	2.6	1.4	1.7
VDFE	Slope	S	2.1	1.9	2.2	2.2	2.1	2.0	1.9	2.4	2.1
VDFE	Slope	N	243	111	132	111	132	59	52	52	80
VDFE	S-Intrcp	M	3.6	2.0	4.9	3.8	3.3	2.6	1.2	5.2	4.7
VDFE	S-Intrcp	S	3.7	3.7	3.3	3.6	3.9	3.6	3.7	3.2	3.3
VDFE	I-Intrcp	M	-1.2	-0.7	-1.7	-1.2	-1.2	-0.9	-0.4	-1.5	-1.8
VDFE	I-Intrcp	S	2.5	2.6	2.4	2.6	2.5	2.7	2.5	2.5	2.4
LDCE	Slope	M	2.0	2.0	2.0	2.0	1.9	1.9	2.1	2.2	1.8
LDCE	Slope	S	1.8	1.9	1.7	1.9	1.7	1.9	1.9	2.0	1.6
LDCE	Slope	N	242	113	129	110	132	59	54	51	78
LDCE	S-Intrcp	M	-0.1	-0.8	0.6	-0.4	0.2	-1.0	-0.6	0.2	0.8
LDCE	S-Intrcp	S	3.1	2.5	3.4	2.8	3.3	2.3	2.8	3.2	3.5
LDCE	I-Intrcp	M	0.5	1.0	0.2	0.8	0.4	1.0	0.9	0.4	.0
LDCE	I-Intrcp	S	1.9	1.4	2.2	1.6	2.1	1.2	1.5	1.8	2.4
LSPE	Slope	M	13.4	10.0	16.3	13.7	13.0	10.4	9.6	17.7	15.4
LSPE	Slope	S	12.3	8.8	14.0	12.4	12.2	8.9	8.7	14.8	13.5
LSPE	Slope	N	240	111	129	105	135	57	54	48	81
LSPE	S-Intrcp	M	-0.5	-3.4	1.9	-3.0	1.4	-4.8	-1.9	-0.8	3.5
LSPE	S-Intrcp	S	18.6	15.0	21.0	17.8	19.1	14.7	15.3	20.9	21.1
LSPE	I-Intrcp	M	0.8	1.1	0.6	1.0	0.7	1.3	0.9	0.7	0.5
LSPE	I-Intrcp	S	1.6	1.5	1.7	1.4	1.8	1.0	1.8	1.7	1.7
SRDE	Slope	M	0.9	0.8	0.9	0.9	0.8	0.8	0.9	1.1	0.8
SRDE	Slope	S	0.6	0.7	0.6	0.7	0.6	0.7	0.6	0.7	0.6
SRDE	Slope	N	243	111	132	110	133	58	53	52	80
SRDE	S-Intrcp	M	-0.1	-0.4	0.2	-0.3	0.1	-0.4	-0.4	-0.1	0.4
SRDE	S-Intrcp	S	1.1	0.8	1.2	0.9	1.1	0.8	0.8	1.0	1.2
SRDE	I-Intrcp	M	0.7	1.1	0.3	0.9	0.5	1.2	1.1	0.6	.0
SRDE	I-Intrcp	S	1.7	1.3	1.8	1.4	1.8	1.3	1.4	1.5	2.0

Table 30

Interactive Reading Assessment System - English:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Comprehension Scales Overall and by Language Category

Scale	Measure	Statistic	Overall	Low Eng	High Eng	Low Span	High Span	Lo Lo	Lo Hi	Hi Lo	Hi Hi
NRCE	Slope	M	2.1	2.1	2.2	2.0	2.2	1.8	2.4	2.3	2.1
NRCE	Slope	S	1.7	1.8	1.6	1.8	1.6	1.9	1.6	1.7	1.5
NRCE	Slope	N	246	113	133	111	135	59	54	52	81
NRCE	S-Intrcp	M	-1.0	-2.0	-0.2	-1.0	-1.0	-1.6	-2.4	-0.4	-0.1
NRCE	S-Intrcp	S	2.9	3.2	2.4	2.9	2.9	3.3	3.0	2.3	2.5
NRCE	I-Intrcp	M	0.9	1.3	0.6	0.9	0.9	1.0	1.6	0.8	0.5
NRCE	I-Intrcp	S	1.3	1.2	1.3	1.2	1.4	1.1	1.2	1.2	1.4
ERCE	Slope	M	2.1	2.1	2.2	2.0	2.3	1.7	2.5	2.2	2.1
ERCE	Slope	S	1.8	1.9	1.7	1.8	1.8	1.9	1.9	1.7	1.7
ERCE	Slope	N	243	111	132	110	133	58	53	52	80
ERCE	S-Intrcp	M	-1.2	-2.3	-0.3	-1.2	-1.3	-1.7	-3.0	-0.6	-0.2
ERCE	S-Intrcp	S	3.6	3.8	3.2	3.1	4.0	3.5	4.0	2.5	3.6
ERCE	I-Intrcp	M	0.8	1.3	0.4	0.8	0.8	1.0	1.6	0.7	0.3
ERCE	I-Intrcp	S	1.8	1.4	2.1	1.5	2.0	1.2	1.5	1.8	2.2
NLCE	Slope	M	1.7	2.0	1.4	1.7	1.6	1.9	2.0	1.4	1.4
NLCE	Slope	S	1.4	1.5	1.3	1.7	1.2	1.8	1.2	1.6	1.2
NLCE	Slope	N	245	113	132	110	135	59	54	51	81
NLCE	S-Intrcp	M	2.0	0.6	3.2	1.9	2.1	0.8	0.4	3.3	3.1
NLCE	S-Intrcp	S	2.9	2.7	2.6	3.0	2.8	2.8	2.6	2.8	2.4
NLCE	I-Intrcp	M	-0.9	0.2	-1.8	-0.7	-1.0	0.2	0.3	-1.8	-1.9
NLCE	I-Intrcp	S	2.4	1.9	2.4	2.3	2.5	1.8	2.1	2.4	2.3
ELCE	Slope	M	1.9	2.3	1.5	1.9	1.9	2.1	2.6	1.9	1.4
ELCE	Slope	S	1.7	1.9	1.5	1.9	1.5	2.0	1.6	1.8	1.3
ELCE	Slope	N	243	111	132	110	133	58	53	52	80
ELCE	S-Intrcp	M	1.1	-0.7	2.6	1.0	1.2	-0.1	-1.4	2.1	3.0
ELCE	S-Intrcp	S	3.6	3.4	2.9	3.3	3.8	3.4	3.3	2.8	3.0
ELCE	I-Intrcp	M	-0.2	0.8	-1.1	.0	-0.4	0.6	1.0	-0.6	-1.4
ELCE	I-Intrcp	S	2.3	1.8	2.3	2.0	2.5	1.7	1.9	2.1	2.4

359

language entry categories (L. Lo, Lo Hi, Hi Lo, and Hi Hi). The number appended to each of these labels is simply the site identification number.

For the overall sample, the student-intercepts for Real Word Decoding, Synthetic Word Decoding, Synthetic Word Spelling, and Sentence Reading are all close to 0 -- in the aggregate, decoding skills are minimal at entry to first grade. For Vocabulary Definition, the student-intercept is at 3.6, almost two grade levels above that expected from the growth track model. The student-intercepts for Listening Comprehension are similarly above expectations, t h not as high as in Definitions, with Narrative at 2.0, and Expository at 1.1. For Reading Comprehension, both the Narrative and Expository scales show student-intercepts approximating -1.0. Thus, under the growth track model, these bilingual students, in the aggregate, enter first grade with English oral language skills which exceed expectation; their decoding skills, as expected, have yet to be developed. Note, however, that the standard deviations are sizeable (for the measures applicable to the growth track, the average is around 3.0), and thus, the aggregate picture is not completely mirrored at the level of the individual student.

The instruction-intercepts provide a similar picture -- those measures carrying a decoding skill component show intercepts at about the level expected by the growth track, while the oral language tasks are associated with negative intercepts. Note that the instruction-intercepts do not exactly correspond to what would be expected given the slope and Y-intercept values. The discrepancy lies in the treatment of the special cases mentioned above, namely, resetting extreme negative outliers to -5.0, and no-growth cases to 0. These treatments must be kept in mind whenever comparisons are made between the three growth measures.

For the aggregated slope measures, Real Word Decoding is seen to show an average rate of 3.2, which is above that expected by the growth track. Similarly, for Synthetic Word Decoding, the aggregate slope value of 2.0, suggests that most students would reach the limits of the testing material by third-grade exit. Thus, coupled with the 0 entry values, decoding of isolated words would not seem to prohibit progress in reading acquisition for these students, as the estimates provide above grade-level expectations throughout the years assessed.

This must be somewhat tempered by the data from the Synthetic Word Spelling task. The slope estimate of 13.4 and the near zero entry estimate, suggest that even by fourth-grade exit, the average student would still be unable to correctly spell 40% of the words contained in the materials.

For each of the oral language tasks (Vocabulary Definition and Narrative and Exposit y Listening Comprehension), the aggregate slope values are just sligh below those expected by the growth track model. Thus, the above expectation entry skills shown in these tasks would be maintained over the years these students were assessed.

Again, in the aggregate, oral language comprehension at both the level of vocabulary and connected text, would not seem to be a barrier to meeting the expectations of : : growth track in reading comprehension. The differences in performance on the vocabulary and connected text tasks are, however, important, and supports the notion that there is much more to understanding a text than simply understanding its constituent words.

As found in Listening Comprehension, the aggregate slope values for Reading Comprehension approximate the value expected from the growth track (2.1 for both narratives and exposition). Given the entry value of about a half grade level below expectation, the reading comprehension of these students is seen to continue to be below grade level, although they are not falling further behind. Finally, the Sentence Reading aggregate slope, .9, coupled with the 0 entry value, indicates that fluency in decoding may present problems in reading connected text. By the end of second grade, the average student would still have a reading rate of less than two syllables per second.

English analyses of variance. Given this aggregate overview, does the description differ for the various entry language groups and sites? To address this question, a series of analyses of variance were conducted on the scale growth measures. Each of the three main summary measures of growth (slope, student-intercept, and instruction-intercept) for each of the nine IRAS scales were subjected to a 2 x 2 x 2 x 5 between-subject analysis, comprising the variables of English entry (low and high), Spanish entry (low and high), gender (boy and girl), and site (0, 1, 2, 3, and 5). Given the difficulty of interpreting three-way interactions, only two-way interactions were analyzed. In the discussion that follows, only those main effects significant at the .05 level are discussed, and only those interactions significant at the .01 level.

The results of the analyses are summarized in Table 31. The left-hand margin is defined by the nine IRAS scales, with the analysis summaries for each of the three growth indices given under each scale. The first columns give the F-ratio and p-values for the main effects and interactions. Given that not all of the variables showed systematic effects, only a subset have been tabled: the main effects of English entry and site, and the interactions of English entry by Spanish entry and English entry by site. The final two columns of the table provide information regarding the amount of variance accounted for by the entered variables, listing (a) the residual sum of squares and (b) the ratio of the explained sum of squares (considering all of the main effects and interactions) to the residual sum of squares. Extracts from the full set of ANOVA summary tables on which Table 31 is based, are presented in Appendix C (nine tables, one for each of the IRAS scales, with three summary tables given within each, corresponding to the three growth indices analyzed).

The summary tables reveal that the most significant variables from the set for the English IRAS scales are those of English entry and

Table 31

Interactive Reading Assessment System - English:
Summary of the Analyses of Variance
on Growth Indices for the Bilingual Sample

Scale - Measure		Variables: F-ratio/p-value				Residual Sum Sqrs	Sum Sqrs Explain/ Residual
		English	Site	Eng x Spn	Eng x Site		
VDC	Slp	5.5/.020	5.5/.001	8.3/.004	1.2/NS	835.2	.25
	IIn	21.3/.001	7.5/.001	2.0/NS	.6/NS	1540.6	.32
	XIn	15.3/.001	3.6/.007	2.0/NS	1.1/NS	412.2	.19
VDF	Slp	8.9/.003	1.0/NS	4.6/NS	1.4/NS	954.7	.12
	IIn	46.8/.001	.5/NS	.7/NS	2.6/NS	2577.3	.32
	XIn	8.7/.004	.5/NS	.5/NS	3.3/NS	1338.5	.15
LDC	Slp	.6/NS	4.1/.003	1.7/NS	2.8/NS	644.5	.18
	IIn	17.9/.001	8.4/.001	.0/NS	1.4/NS	1718.5	.32
	XIn	15.2/.001	5.2/.001	.1/NS	1.4/NS	667.2	.28
LSP	Slp	11.6/.001	3.8/.006	1.6/NS	1.3/NS	29787.3	.21
	IIn	7.8/.006	8.2/.001	3.0/NS	1.8/NS	64679.7	.28
	XIn	8.6/.004	4.5/.002	5.3/NS	2.8/NS	498.1	.24
SRD	Slp	.0/NS	14.0/.001	3.3/NS	.3/NS	76.8	.33
	IIn	31.0/.001	8.3/.001	1.1/NS	1.5/NS	194.9	.41
	XIn	20.3/.001	4.6/.001	.8/NS	.9/NS	503.4	.33
NRC	Slp	.0/NS	7.6/.001	1.4/NS	.7/NS	582.7	.18
	IIn	26.7/.001	12.8/.001	.0/NS	3.0/NS	1461.0	.44
	XIn	19.2/.001	6.5/.001	2.5/NS	1.7/NS	318.6	.30
ERC	Slp	.0/NS	3.4/.009	1.2/NS	.9/NS	672.3	.15
	IIn	19.5/.001	6.7/.001	.3/NS	2.3/NS	2367.5	.35
	XIn	13.9/.001	6.2/.001	.5/NS	1.7/NS	636.6	.28
NLC	Slp	9.9/.002	1.1/NS	[illegible].2/NS	1.1/NS	454.6	.11
	IIn	52.6/.001	1.8/NS	.7/NS	1.0/NS	1518	.35
	XIn	46.0/.001	3.4/.010	1.7/NS	1.2/NS	1030	.37
ELC	Slp	16.3/.001	4.4/.002	3.6/NS	.4/NS	599.4	.19
	IIn	65.3/.001	3.5/.009	.2/NS	1.1/NS	2156.2	.42
	XIn	42.1/.001	.9/NS	.2/NS	2.1/NS	960.7	.31

Note: Significance levels are .05 for main effects, .01 for interactions.

362

BEST COPY AVAILABLE

site. Thus, the discussion which follows will concentrate on the differences expressed in these variables, treating English entry first.

English entry is found to have significant effects on all three of the growth measures with the exception of the slope indices for four scales (namely, Synthetic Word Decoding, Sentence Reading, and Narrative and Expository Reading Comprehension). Figures 13 through 21 display the growth functions for the two groups, along with the overall function, for each of the nine scales (the data are taken from Tables 29 and 30; these, and all subsequent plots, are based on the slope and y-intercept values).

The figures show that the high English entry group consistently has higher student-intercept values and lower instruction-intercept values than the low English entry group. For the significant slope differences, the high English entry group grows faster in the decoding component tasks (Real Word Decoding and Synthetic Word Spelling), but the low English group has larger growth rates in the oral language tasks (Vocabulary Definitions, and Narrative and Expository Listening Comprehension).

Recall that the aggregate picture for decoding skills was entry at 0 with growth above expectation. Here the high English group enters above expectation (by about a half IRAS level), while the low English entry group enters below expectation (by about three-quarter IRAS levels); and where growth differs, while above the expected values for both groups, it is greater for the high English entry group (by about three-quarter IRAS levels).

For the oral language task of Vocabulary Definition, the high English entry group enters at about the fifth IRAS level, showing a three level advantage over the low English entry group. However, growth for the latter is slightly above the expected rate (at 2.2) and about three-quarters IRAS levels above the high English entry group -- thus, a convergence of skill would be expected by fourth-grade exit (see Figure 14). For the other oral language tasks of Listening Comprehension, similar differences in growth rates are found, and also in entry values (although, the entry values are about two levels lower than those for Vocabulary Definition).

For the Reading Comprehension scales, growth rates do not differ between the English entry groups -- both proceed at a rate approximating that expected from the growth track model, with narrative rates being slightly higher (2.2), and exposition slightly lower (1.8). The intercept differences, however, are significant -- the high English entry group enters at about the expected level, while the low English entry group is two levels below expectation. Thus, for reading comprehension, the high English entry group begins first grade with a grade-level advantage over the low English entry group, but growth is then at a "year for year" rate for each. For Sentence Reading, the two groups do not differ on the slope measures, but the entry intercept is below expectations for the low English entry group (at -.4) and above the expected level for the high English entry group (at .2).

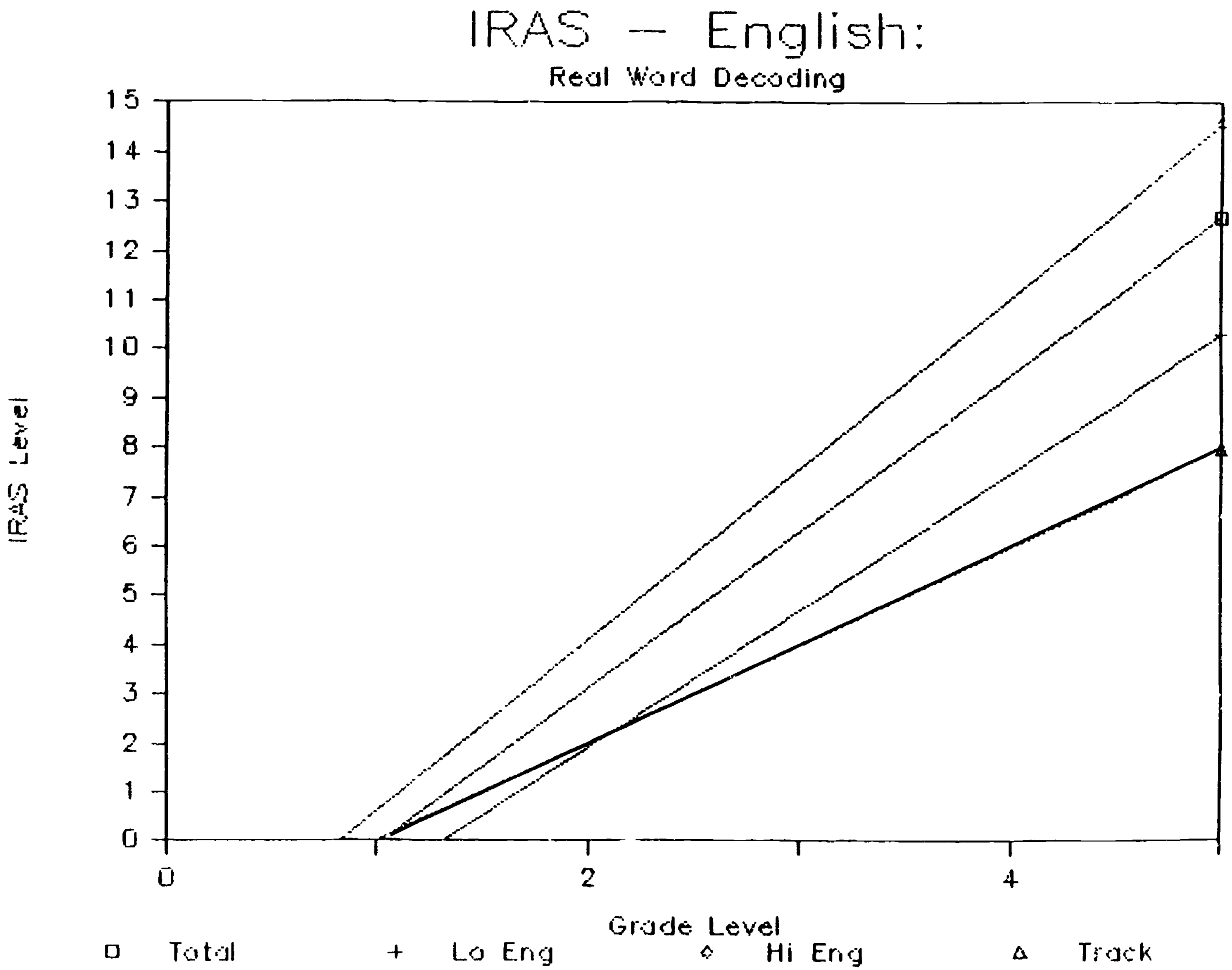


Figure 13. IRAS-E Real Word Decoding growth for the entire bilingual sample and by English entry category.

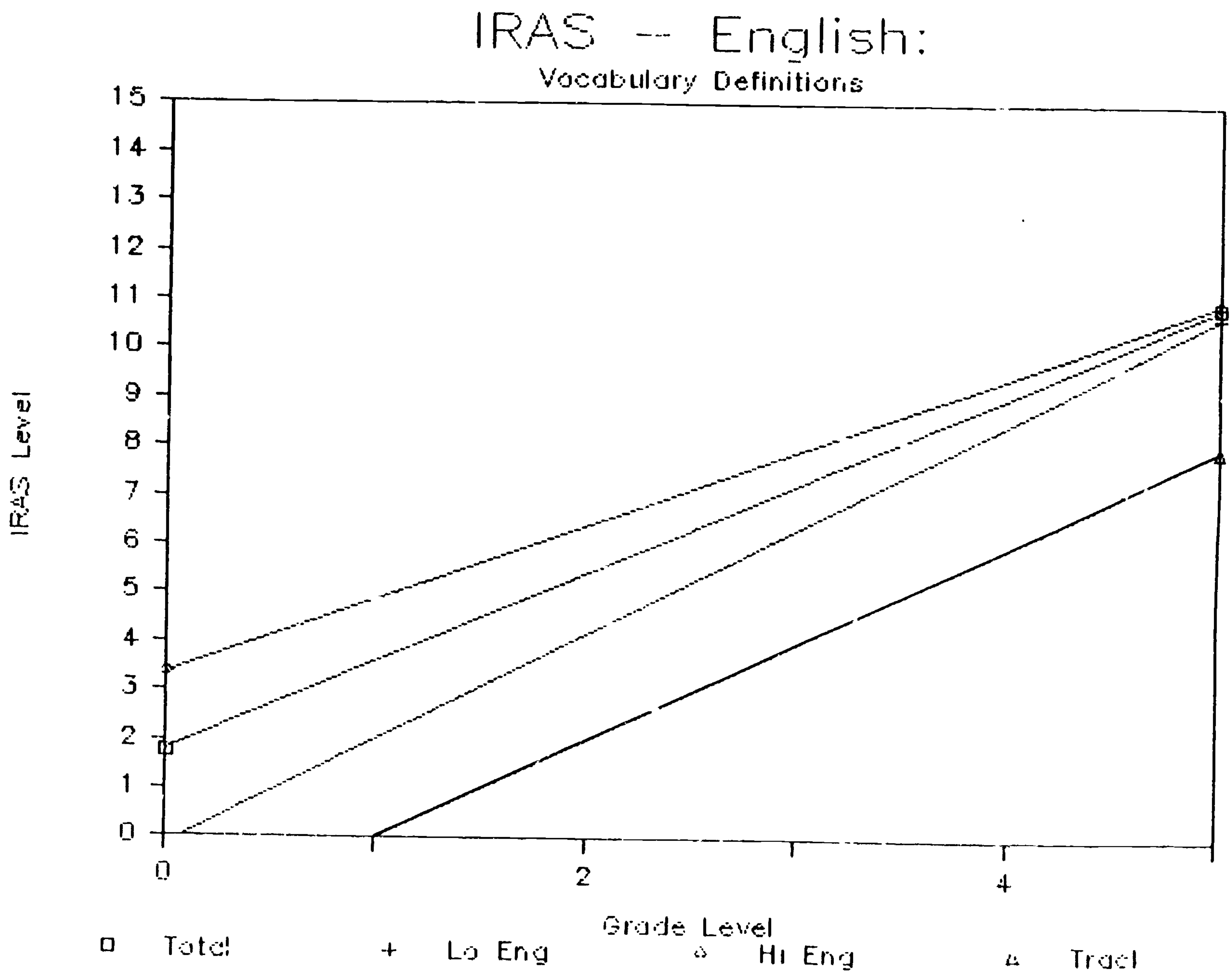


Figure 14. IRAS-E Vocabulary Definition Growth for the entire bilingual sample and by English entry category.

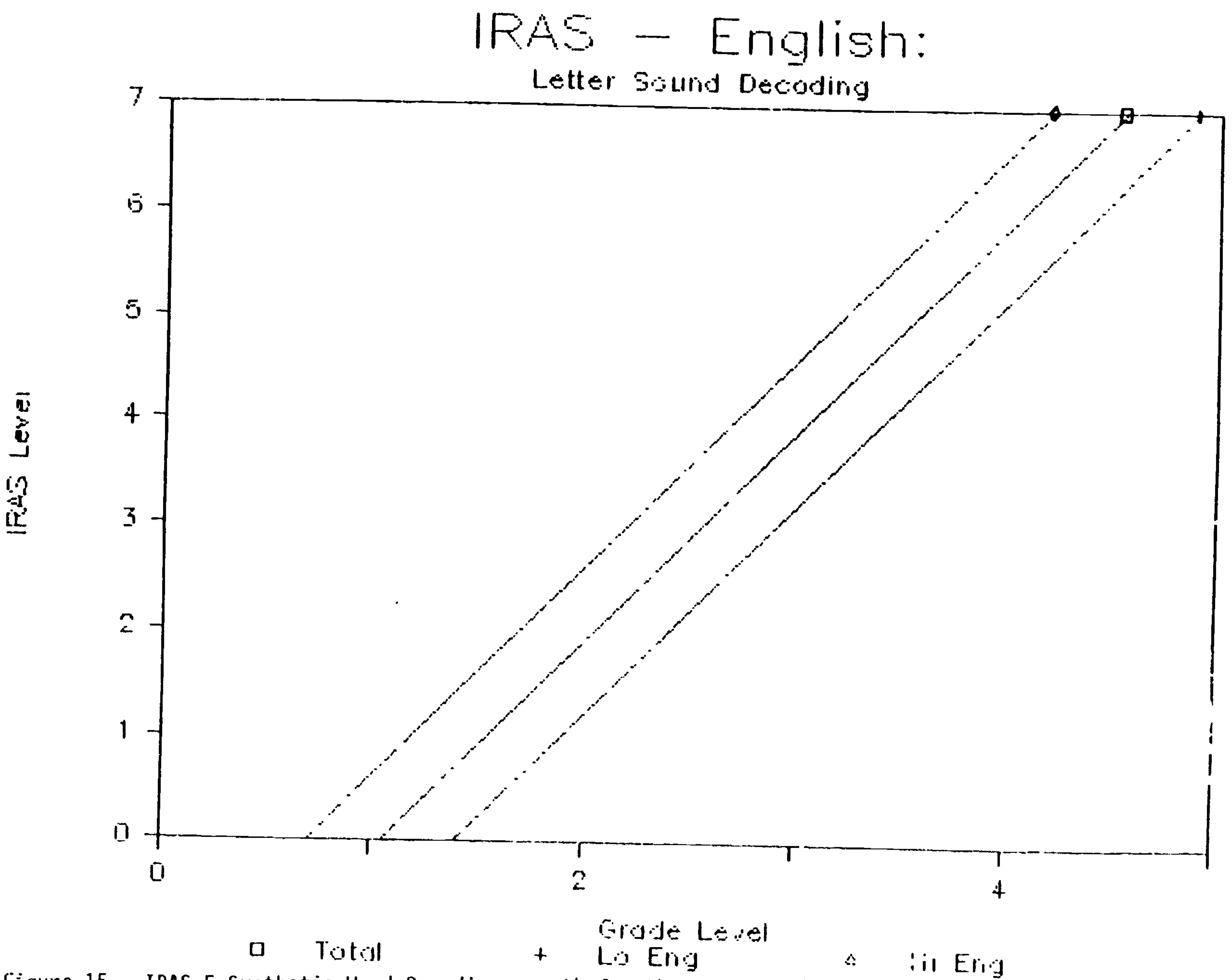


Figure 15. IRAS-E Synthetic Word Decoding growth for the entire bilingual sample and by English entry category.

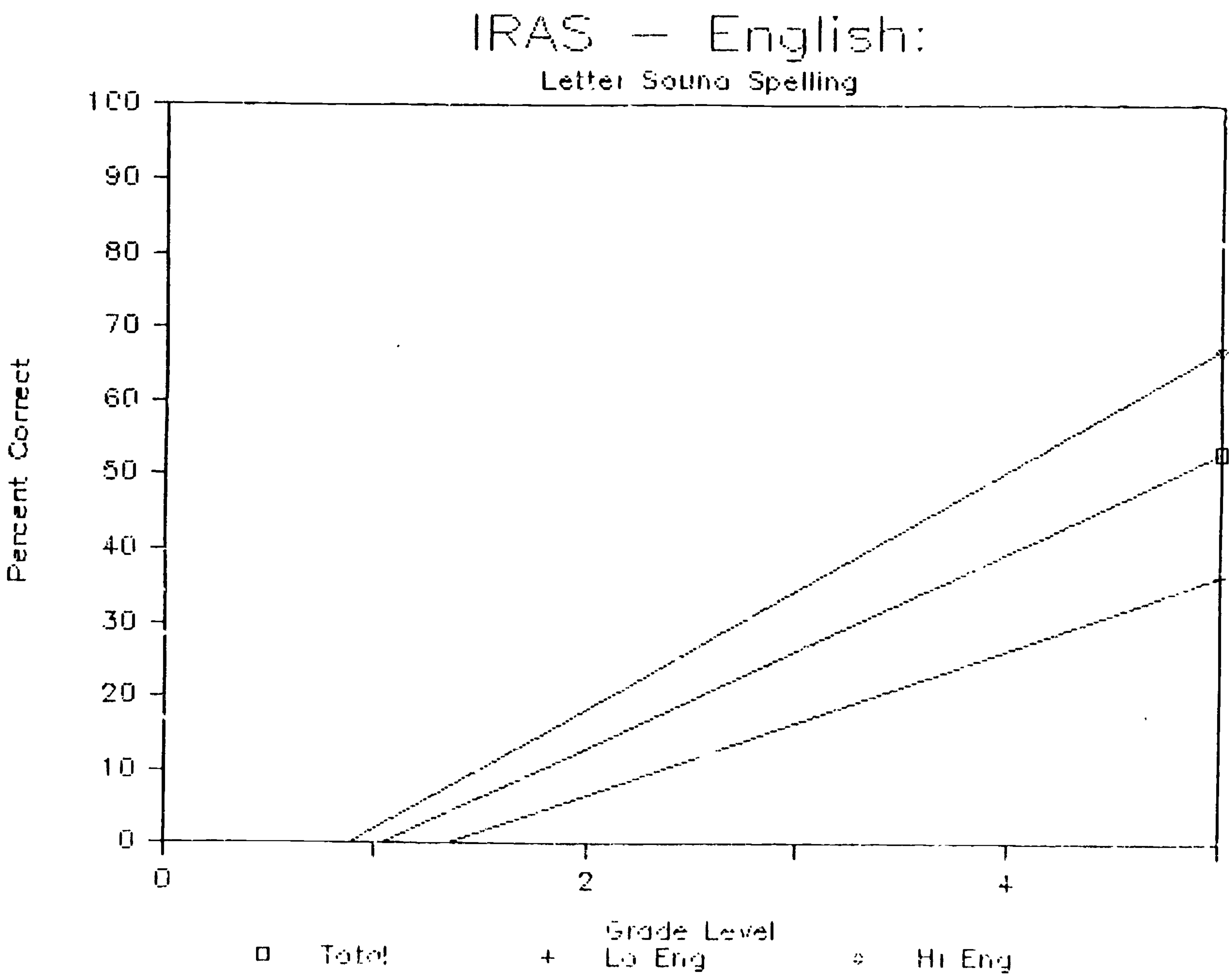


Figure 16. IRAS-E Synthetic Word Spelling growth for the entire bilingual sample and by English entry category.

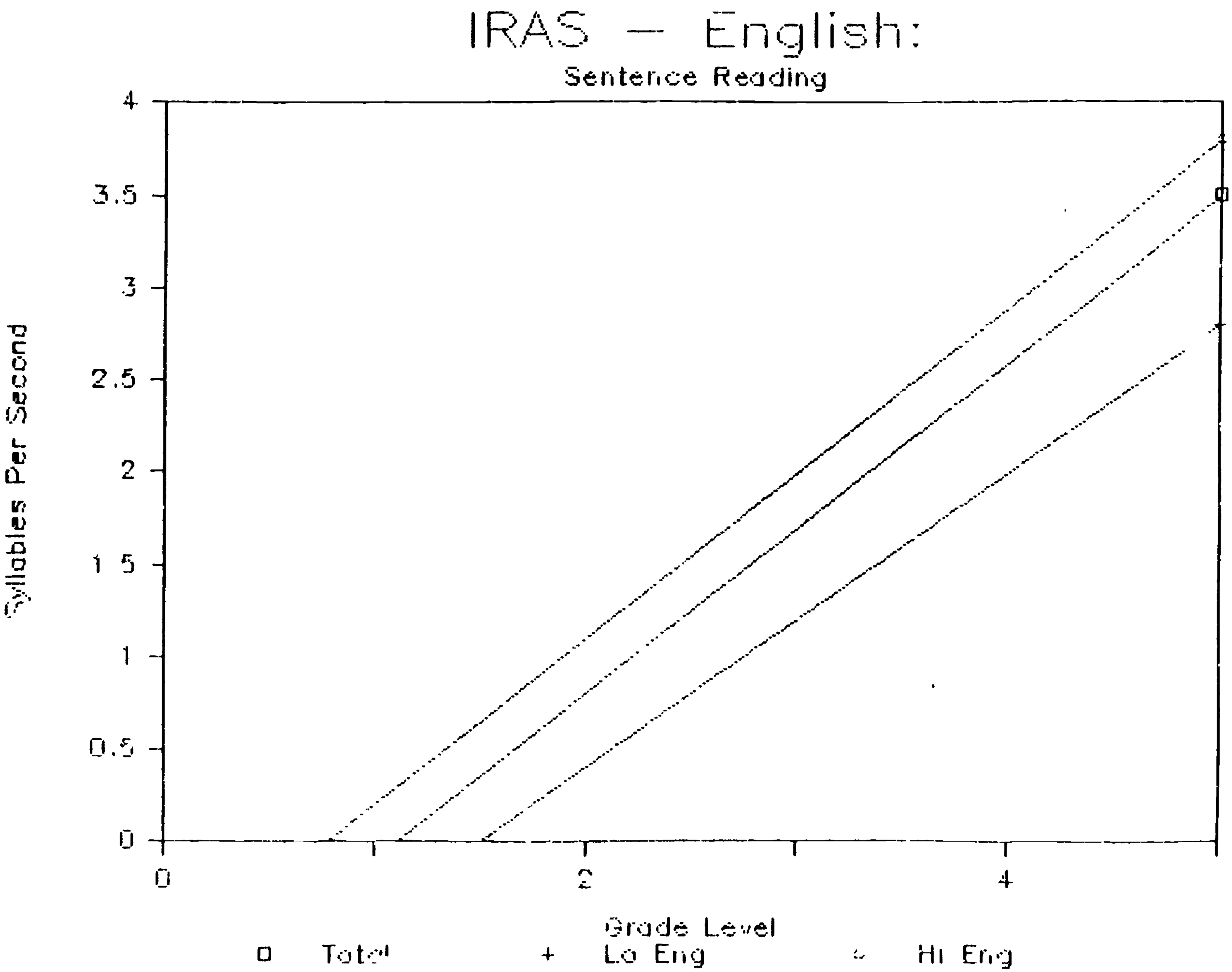


Figure 17. IRAS-E Sentence Reading Growth for the entire bilingual sample and by English entry category.

87

372

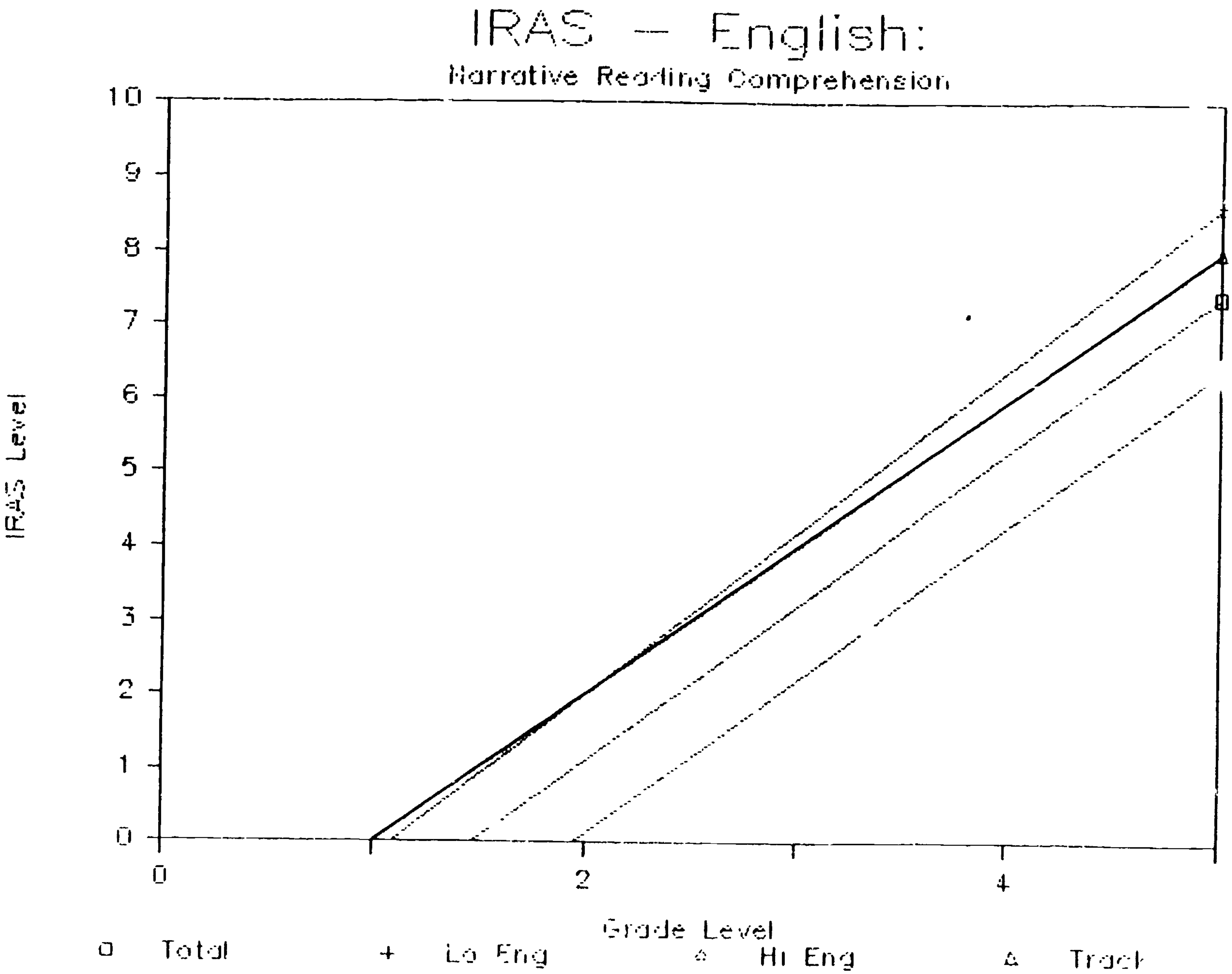


Figure 18. IRAS-E Narrative Reading Comprehension growth for the entire bilingual sample and by English entry category.

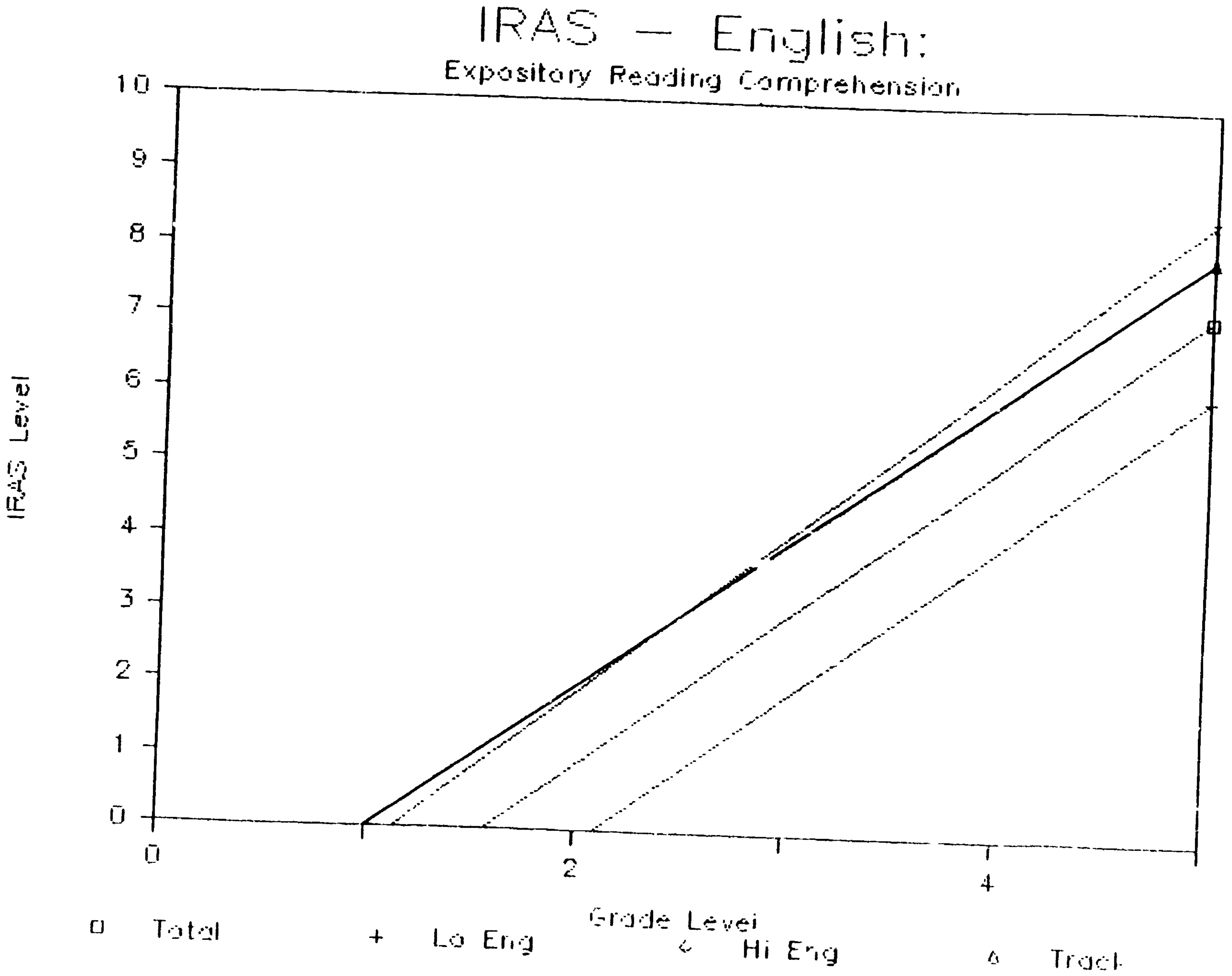


Figure 19. IRAS-E Expository Reading Comprehension growth for the entire bilingual sample and by English entry category.

90

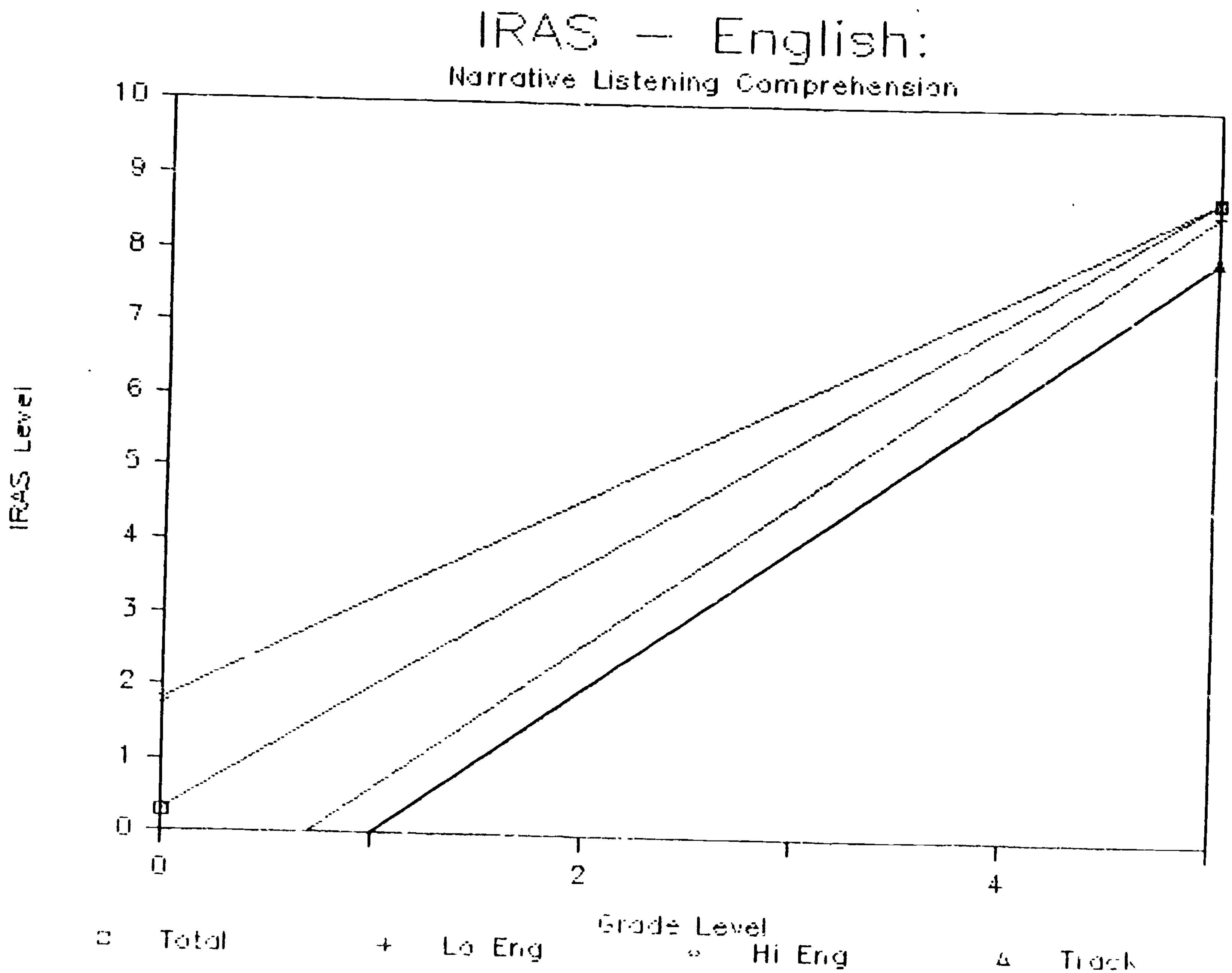


Figure 20. IRAS-E Narrative Listening Comprehension growth for the entire bilingual sample and by English entry category.

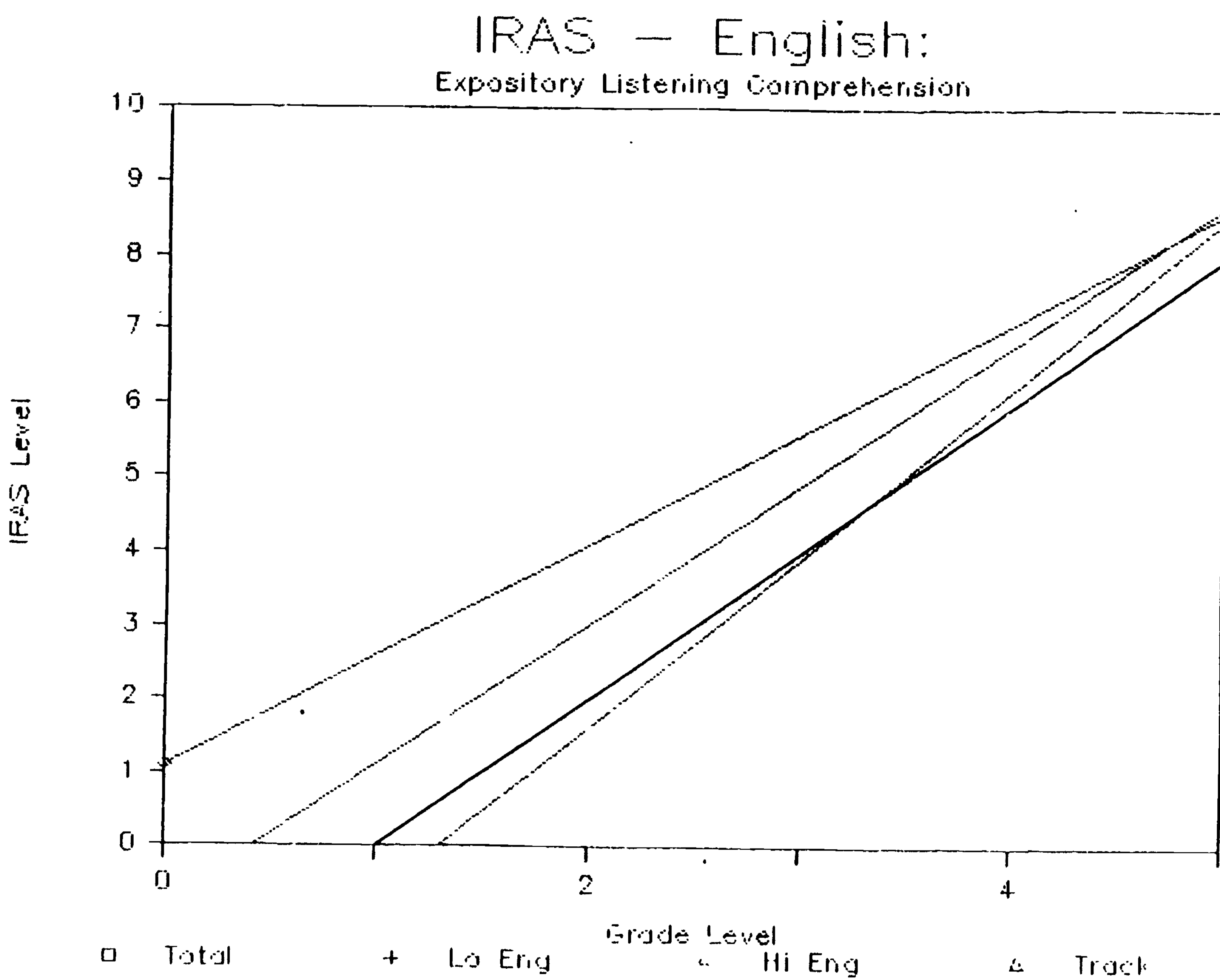


Figure 21. IRAS-E Expository Listening Comprehension growth for the entire bilingual sample and by English entry category.

Overall, the picture for English entry differences appears to be that students who come to school with lower English skills show greater growth in oral language capacities, and thus, show a convergence in such skills with those students who entered with higher English skills. However, the high English entry students are better able to profit from decoding instruction in a way that their initial advantage in decoding continues to expand.

Interestingly, Spanish entry is found to have significant effects on the slope indices (but not the intercepts) of Narrative and Expository Reading Comprehension (see Appendix C). For both tasks, the growth rates are greater (by about .3 IRAS levels) for students with greater Spanish entry skills. Looking at the English reading comprehension slopes for the combined English and Spanish entry groups (Table 30), the growth rate is highest for those students with low English skill and high Spanish skill, and lowest for those students with low skills in both languages. This allows an interpretation of transferability: although the growth rates for English listening comprehension do not differ for these two groups, it seems that relatively higher skills in Spanish at school entry promotes the growth of English reading comprehension.

Site is the only remaining variable found to have widespread effects, covering all of the IRAS scales (except Vocabulary Definition), and all of the growth measures (except the instruction-intercept in Expository Listening Comprehension and the slope and student-intercept in Narrative Listening Comprehension). Figures 22 through 30 display the growth functions for the five sites for each of the nine scales (the data are taken from Appendix B).

Before turning to a discussion of the site differences, remember, as cautioned earlier, that the reliabilities of the estimates of slope and intercept are greater for those cohorts with greater numbers of data points. Thus, when considering site differences, the reliability of the estimates will be greatest for the border sites (Sites 0, 1, and 2), each containing three to four data points, and least for the non-border sites (Sites 3 and 5), where each estimate is based on only two points.

The differences attributable to site generally follow the pattern of high student-intercepts (low instruction-intercepts) coupled with low growth rates, and low student-intercepts (high instruction-intercepts) coupled with high growth rates. In describing the site differences each of the scales will be discussed individually, treating those dealing with formal language (Vocabulary Definition, and Narrative and Expository Listening Comprehension), decoding (Real Word Decoding, Synthetic Word Decoding, and Synthetic Word Spelling), and reading (Sentence Reading, and Narrative and Expository Reading Comprehension) in turn.

Considering the for . language tasks, no differences in the growth indices for site were found in Vocabulary Definition, and, as can be seen in Figure 23, all sites show performance levels above the

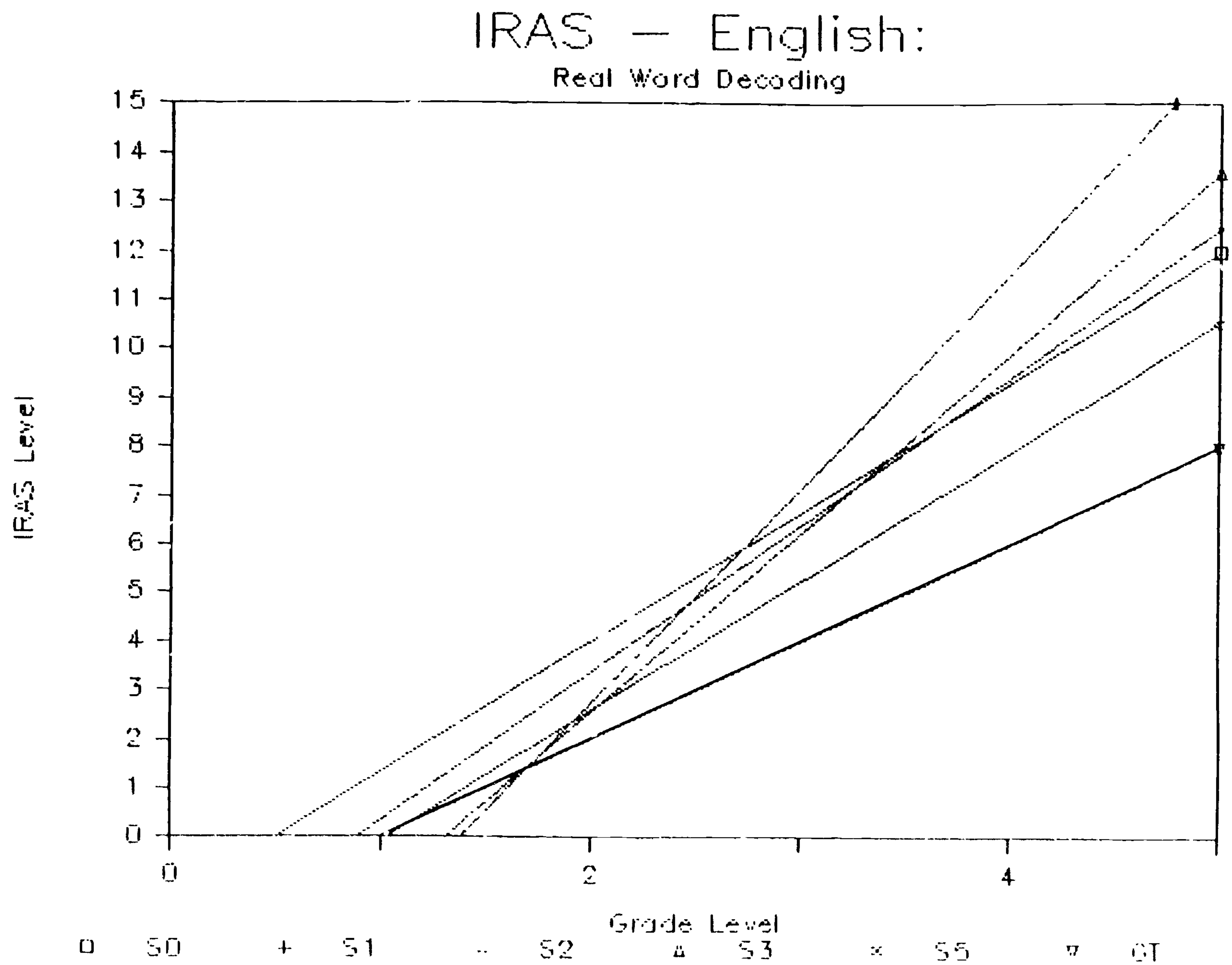


Figure 22. IRAS-E Real Word Decoding growth for each site.

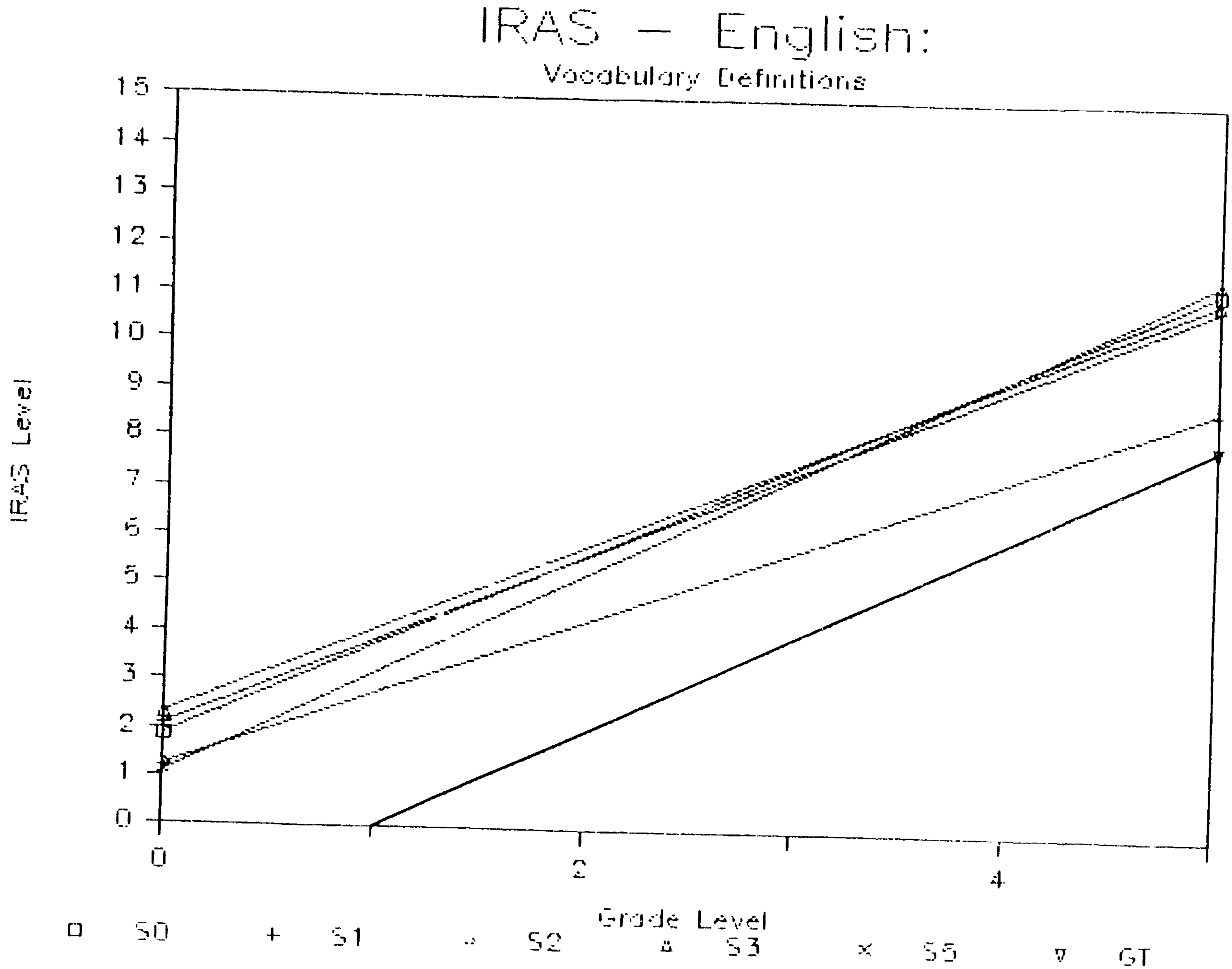


Figure 23. IRAS-E Vocabulary Definition growth for each site.

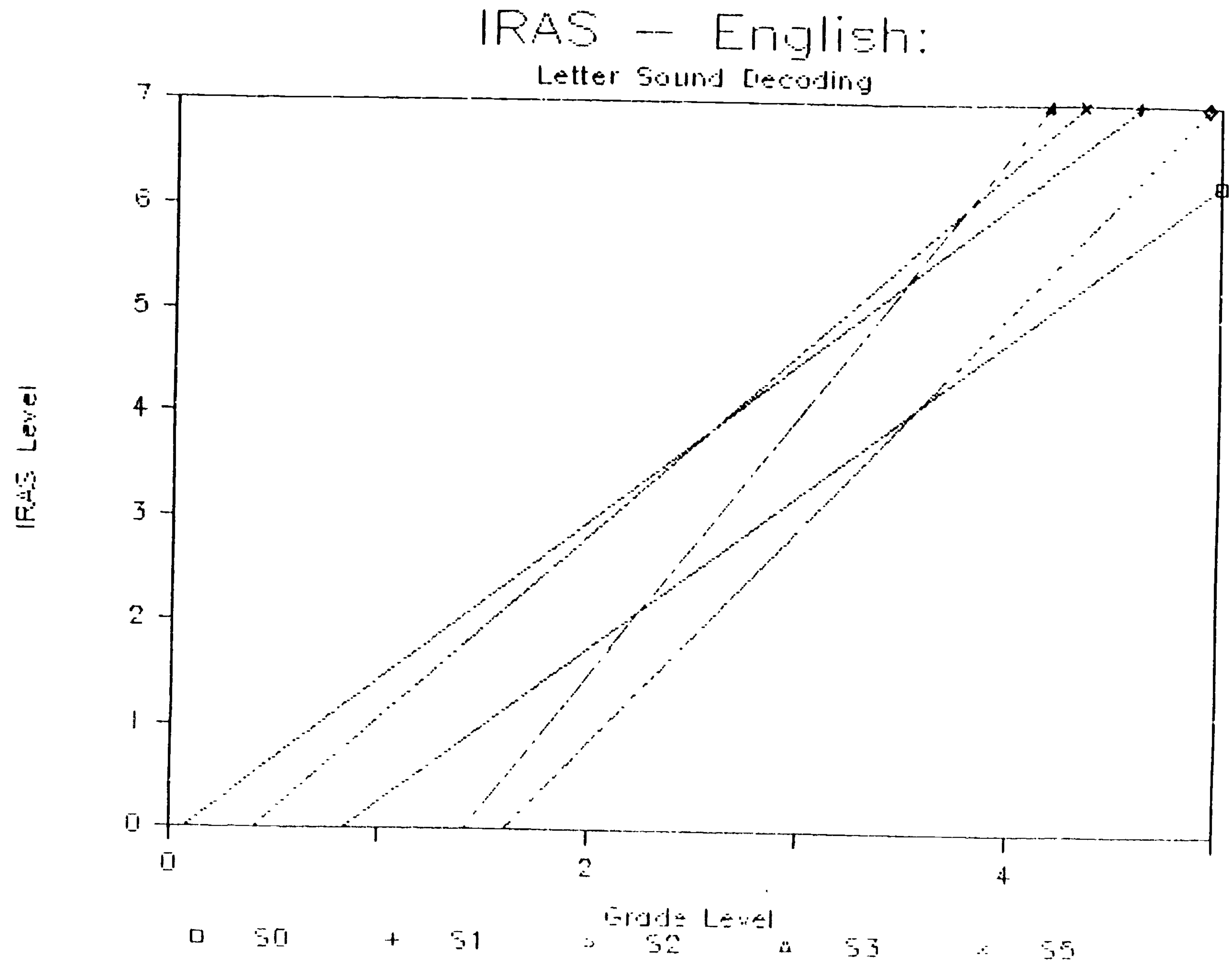


Figure 24. IRAS-E Synthetic Word Decoding growth for each site.

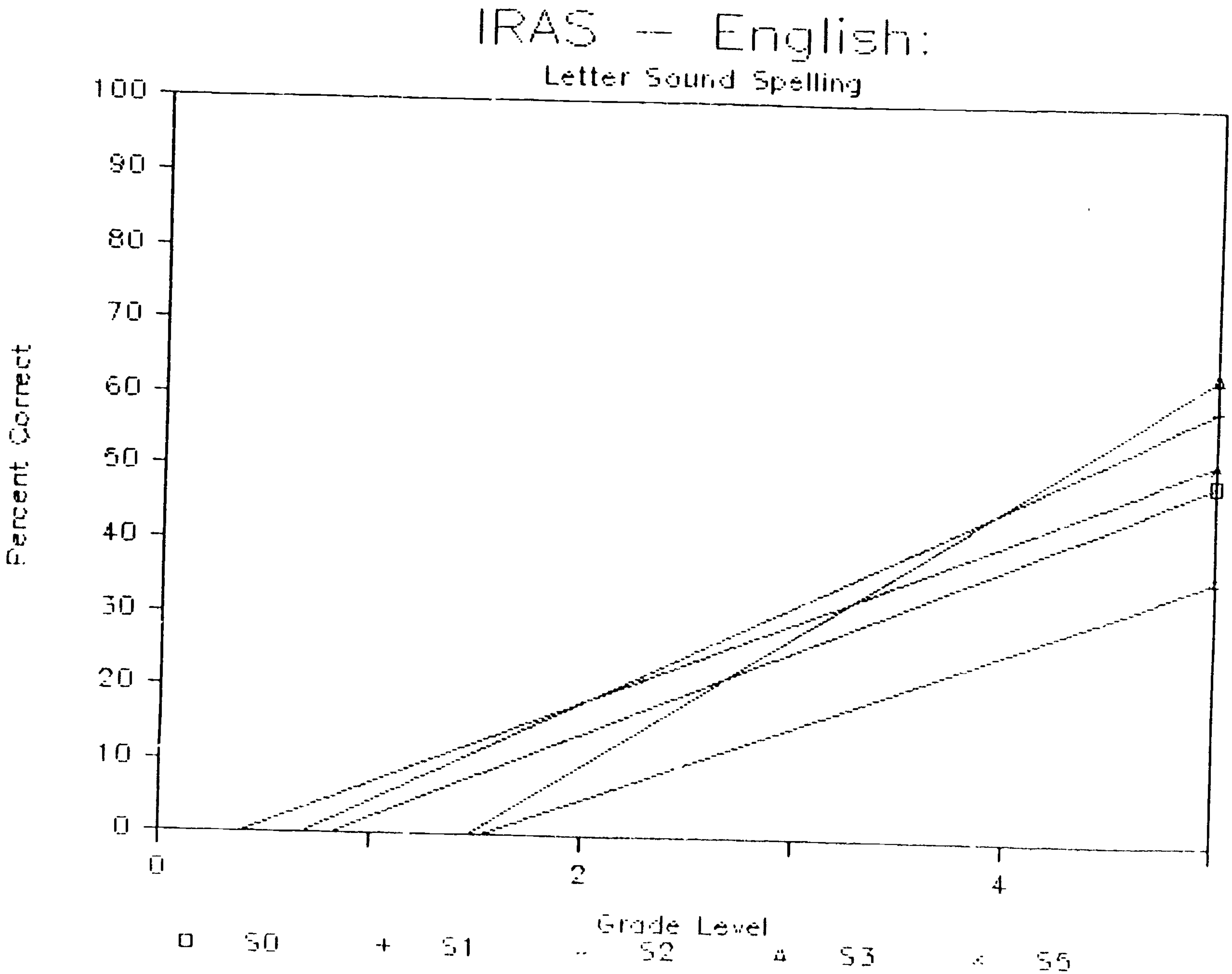


Figure 25. IRAS-E Synthetic Word Spelling growth for each site.

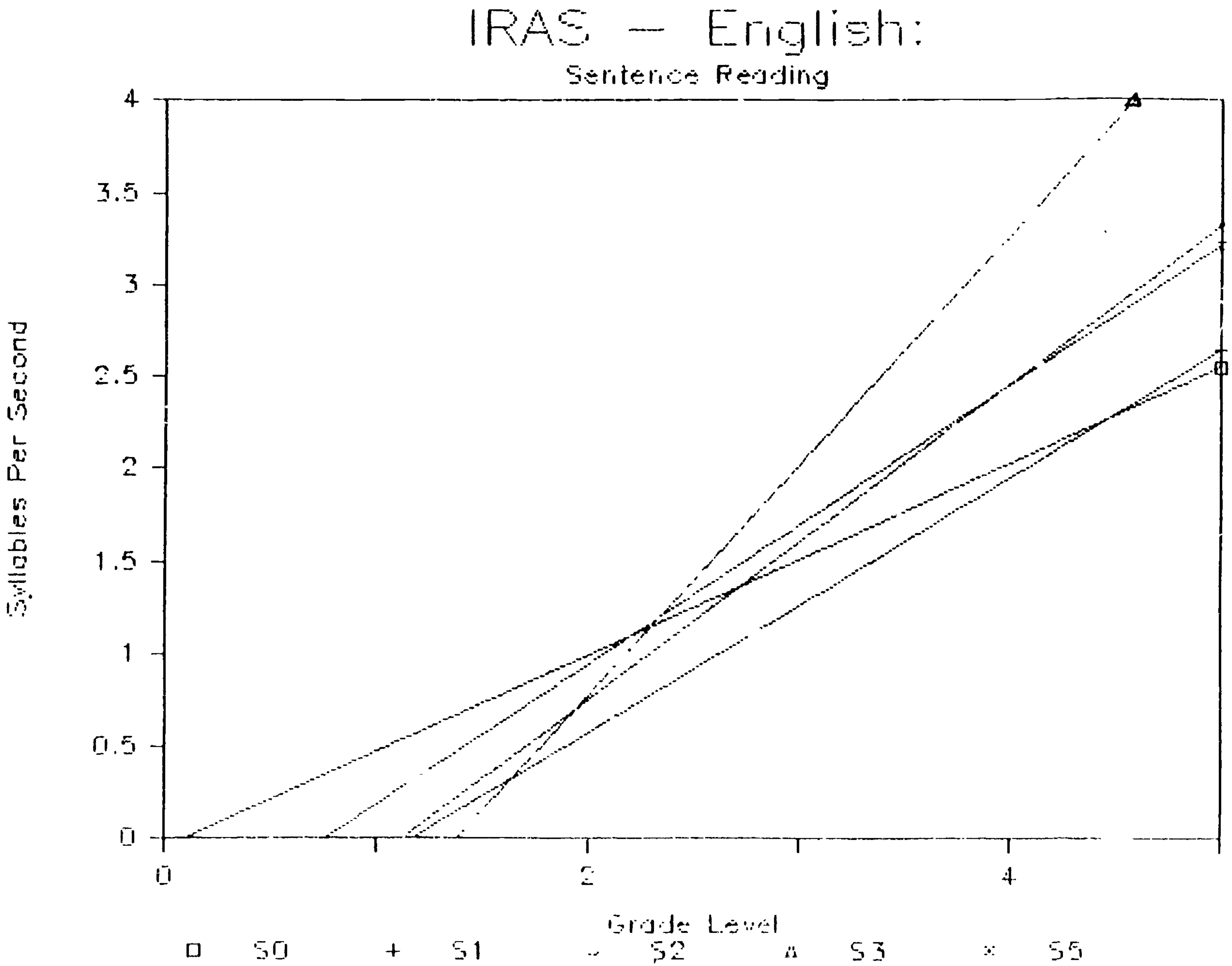


Figure 26. IRAS-E Sentence Reading growth for each site.

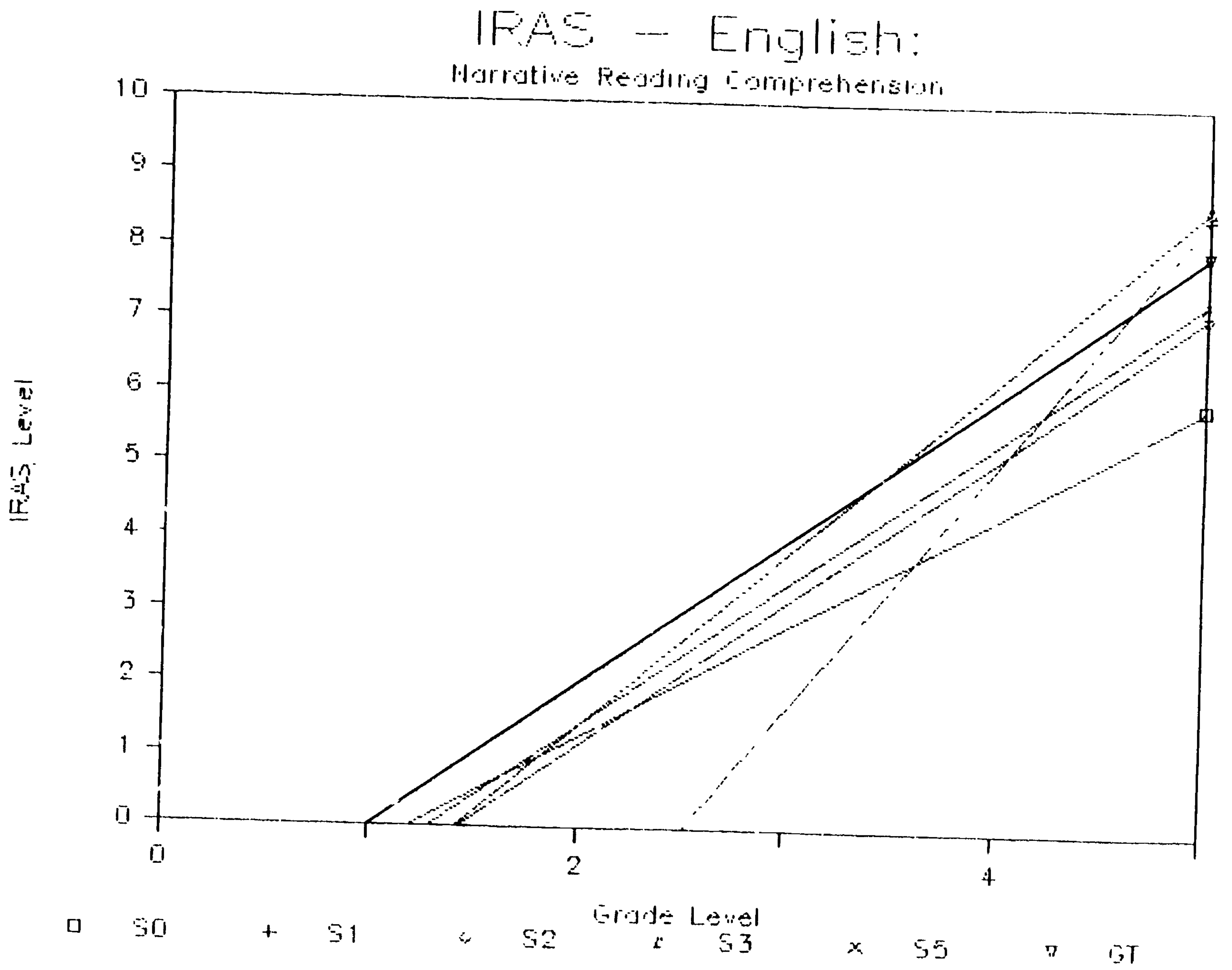


Figure 27. IRAS-E Narrative Reading Comprehension growth for each site.

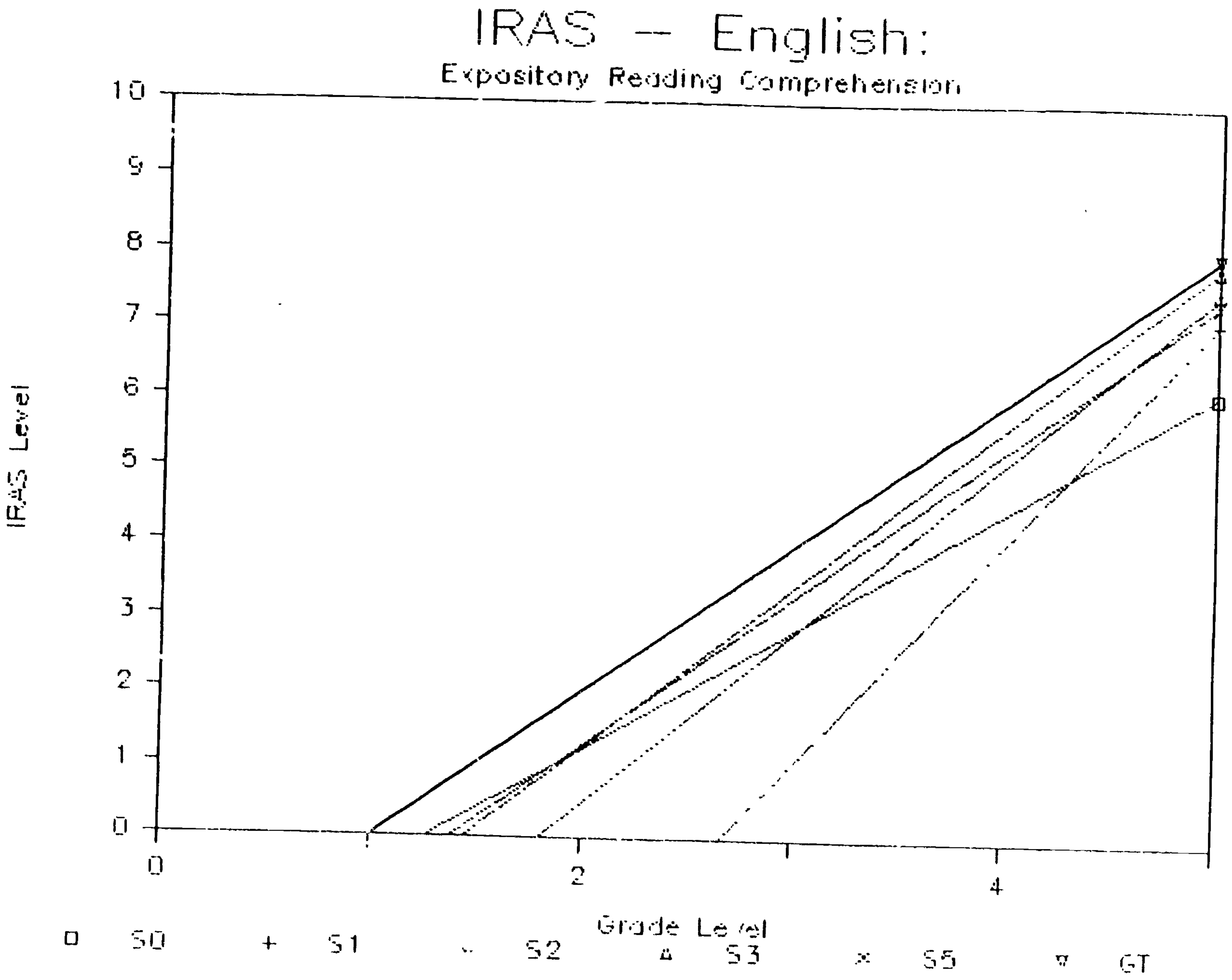


Figure 28. IRAS-E Expository Reading Comprehension growth for each site.

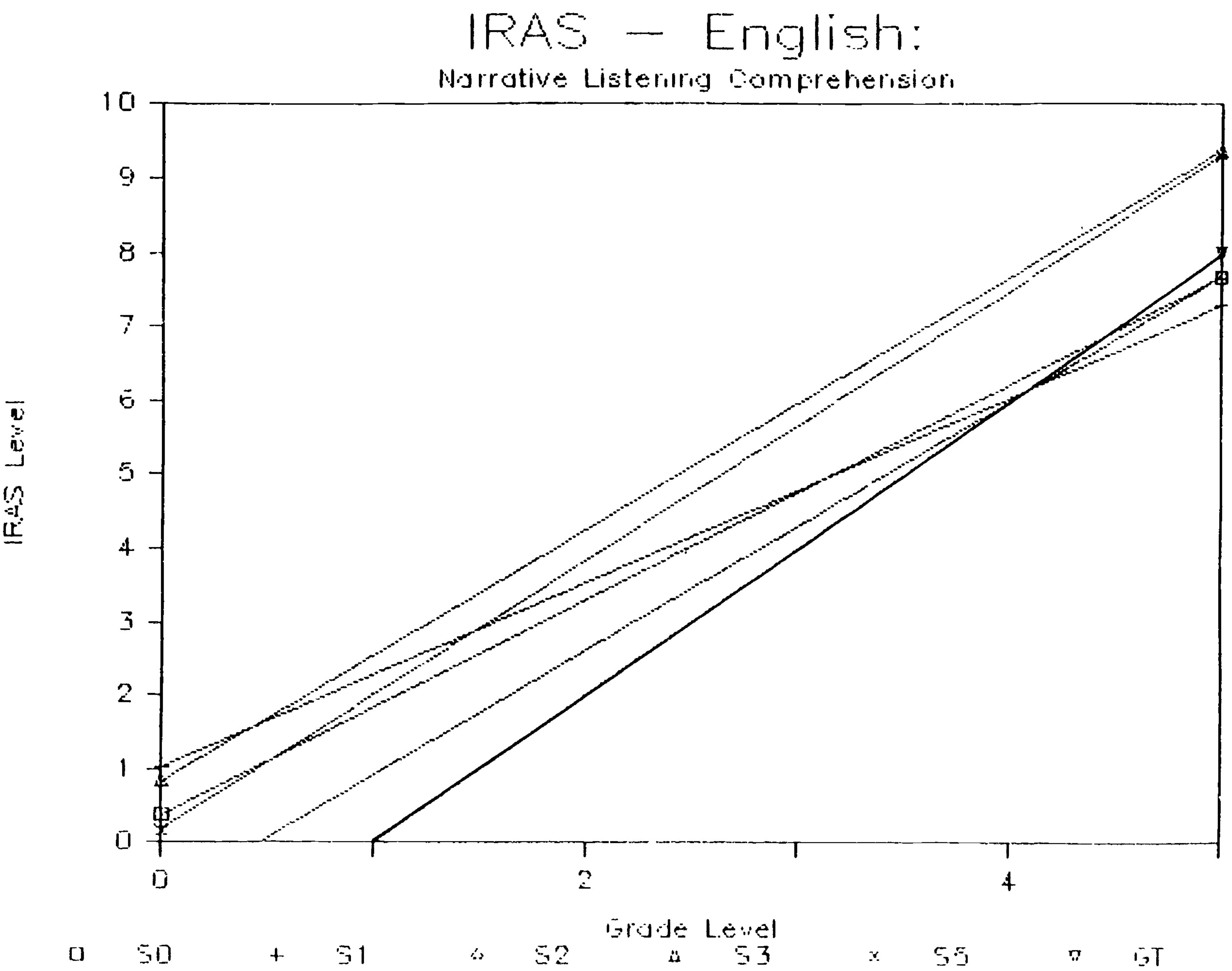


Figure 29. IRAS-E Narrative Listening Comprehension growth for each site.

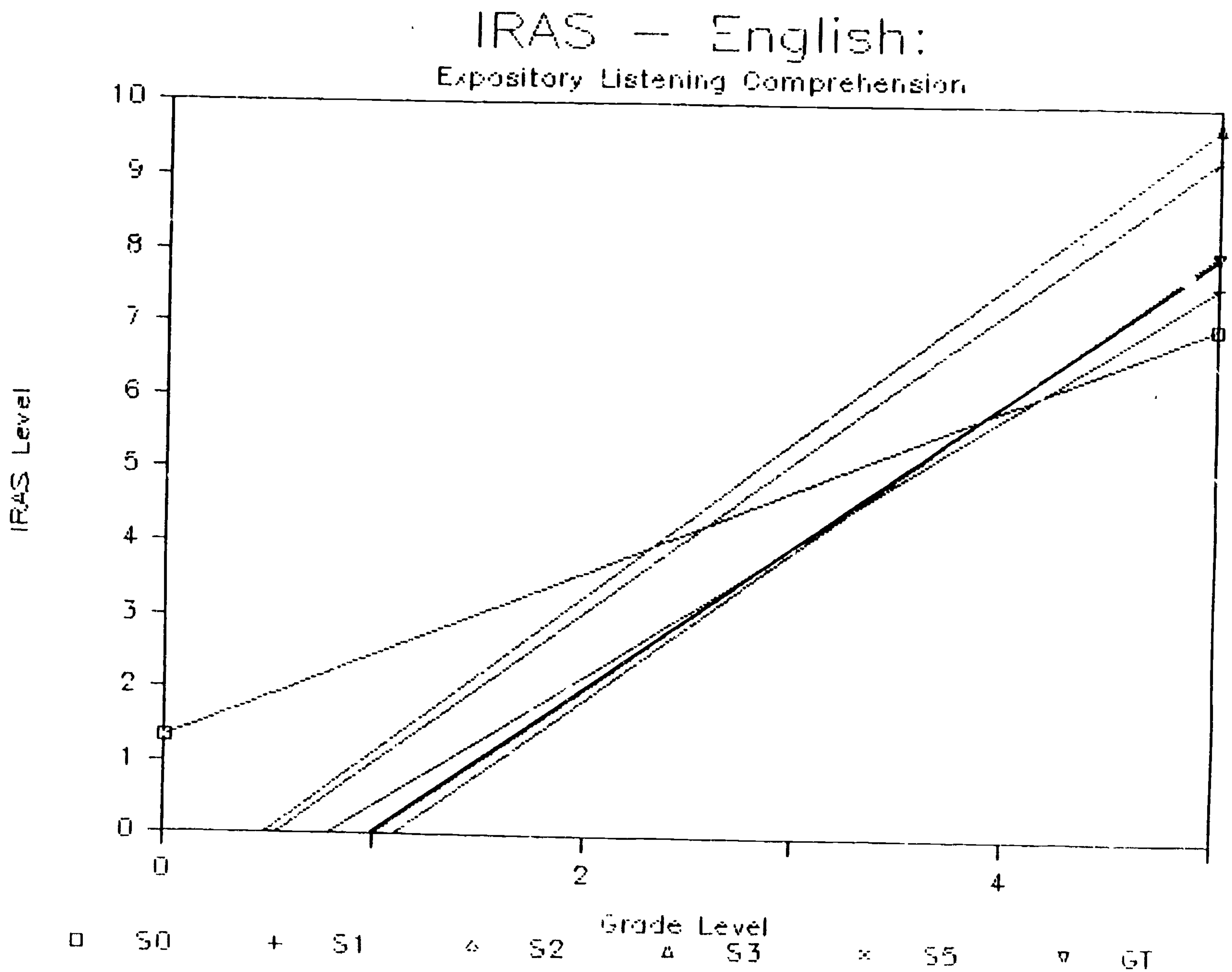


Figure 30. IRAS-E Expository Listening Comprehension growth for each site.

growth track model with growth rates slightly less than that expected -- growth at Site 2 appears to be slightly lower, though the difference is not statistically significant. For Narrative Listening Comprehension, displayed in Figure 29, the growth rates are somewhat smaller than expected, and all student-intercepts are above the model expectations. Only the instruction-intercepts differ by site, most strongly influenced by the value for Site 2 which is substantially greater than those for the other sites. For Expository Listening Comprehension (Figure 30), growth follows that of the growth track for all sites except Site 0 where growth is at a lower rate than that predicted, and begins from a substantially greater student-intercept (about 2.5 IRAS levels). Additionally, the student-intercepts at Sites 3 and 5 are high at about a level above the growth track.

Thus, for the formal language tasks, performance is generally the same across sites with two notable exceptions. First, Site 2 shows a relative delay in the onset of effective instruction in Narrative Listening Comprehension, and Site 0 shows a relatively high student-intercept coupled with relatively slow growth in Expository Listening Comprehension.

For Real Word Decoding (Figure 22), all sites show growth rates which exceed those predicted from the growth track model. The highest rates are at Sites 1 and 3 (with the greatest instruction-intercepts) and the lowest rates are at Sites 0, 2, and 5 (with low instruction-intercepts). For Synthetic Word Decoding (Figure 24), all sites (except Site 0) are expected to reach the limits of the test by fourth-grade exit. Sites 0, 1 and 5 show comparable growth rates; Site 0 has a student-intercept of about zer , and Sites 1 and 5, about one. The greatest growth is shown in Sites 2 and 3, both with comparable high instruction-intercepts. In Synthetic Word Spelling (Figure 25), no site, in the aggregate, is projected to be able to correctly spell more than 60% of the items at fourth-grade exit. The growth rates for all sites are quite similar except that for Site 3, which is higher (recall that the data from this site are based on only two data points and thus, the estimate is less trustworthy). Site 2 shows the least attainment of skill in this task.

Thus, for decoding, the site differences can be characterized as follows. Sites 0 and 5 show instruction-intercepts between 0 and 1, suggesting that effective instruction in decoding began during their kindergarten programs, whereas the instruction-intercepts for Sites 2 and 3 indicate delayed instruction to about the middle of first-grade. Site 1 shows delayed instruction in Real Word Decoding, but early instruction in dealing with synthetic words. The growth rates for Site 3 are generally greater than those for the other sites, which are comparable, although Site 2 shows high growth in Synthetic Word Decoding, and Site 1 reveals high growth in Real Word Decoding.

For the reading tasks, in Sentence Reading (Figure 26), Site 3 shows considerable growth and a low student-intercept (high instruction-intercept); Site 0 shows a low growth rate and a large student-intercept (low instruction-intercept); and the remaining sites show comparable growth patterns which are between these extremes. The growth patterns for individual sites are comparable for Narrative and Expository Reading Comprehension, with performance on the latter slightly lower than that on the former. Performance is below the growth track, with Site 1 showing considerable growth coupled with a substantial delay in the onset of effective instruction. The remaining sites show similar instruction-intercepts, at about the middle of first grade, where growth is at the level expected by the growth track model for Sites 2 and 5, below expectation for Site 0, and above expectation for Site 3.

Thus for reading skill, Site 1 shows a delay of effective instruction, but a high rate of growth once such instruction begins. Site 0, on the other hand, shows relatively early effective instruction, but a relatively low growth rate. Sites 2, 3, and 5 show similar growth patterns across the tasks, although Site 3 has a much higher growth rate in Sentence Reading.

Overall, the differences between sites over the English IRAS assessments can be briefly characterized as follows. First, across each of the areas, Site 0 students generally seem to have encountered relatively early effective instruction, but show relatively low growth over the ensuing years. Site 1 shows slow growth in formal language skill and a delay in effective instruction in English decoding (for real words, but not for synthetic ones) and reading, but high growth once such instruction is encountered. Site 2 shows growth in formal language as expected from the growth track model; for decoding skill, a delay in effective instruction is found, and growth is low (except in Synthetic Word Decoding, where it is quite high). For Sites 3 and 5, formal language skill is high at entry and growth follows expectations. For the decoding and reading tasks, Site 5 closely follows the average over sites, but Site 3 shows delayed instruction and high subsequent growth.

Spanish overall performance. The Spanish IRAS descriptive statistics for the bilingual sample are presented in Tables 32 and 33, for the comprehension and non-comprehension scales, respectively. The formats of these tables are identical to those used for the English data (see the description above). The Spanish IRAS data for the individual sites (Sites 0, 1, 2, 3, and 5) are presented in Appendix D, with the first table of each site-pair providing the data for the non-comprehension scales, the second, the data for the comprehension scales.

For the overall sample, the student-intercepts for decoding (Real Word Decoding, Synthetic Word Decoding, and Synthetic Word Spelling) are all close to 0 -- as found in the English IRAS decoding scales, in the aggregate, the sample begins first grade with little skill in decoding. For Vocabulary Definition, the student-intercept is almost three IRAS levels above that expected by the growth track model. The student-intercepts for Listening Comprehension are similarly above expectation, with narratives one level below that for Vocabulary Definition, and exposition, two levels lower. For the reading scales, the student-intercepts are at zero for Sentence Reading, and about a

Table 32

Interactive Reading Assessment System - Spanish:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Non-comprehension Scales Overall and by Language Category

Scale	Measure	Statistic	Overall	Low Eng	High Eng	Low Span	High Span	Lo Lo	Lo Hi	Hi Lo	Hi Hi
VDCS	Slope	M	2.2	2.4	2.0	1.4	2.8	1.6	3.2	1.1	2.6
VDCS	Slope	S	2.8	3.1	2.5	2.5	2.9	2.8	3.2	2.2	2.6
VDCS	Slope	N	247	112	135	111	136	58	54	53	82
VDCS	S-Intrcp	M	0.3	0.1	0.4	-0.3	0.8	-0.7	1.0	0.0	0.6
VDCS	S-Intrcp	S	4.1	4.4	3.9	2.8	4.9	2.7	5.6	3.0	4.5
VDCS	I-Intrcp	M	0.6	0.7	0.5	0.7	0.5	0.9	0.5	0.6	0.5
VDCS	I-Intrcp	S	1.6	1.6	1.7	1.2	1.9	1.2	1.9	1.2	1.9
VDFS	Slope	M	1.8	1.8	1.7	1.7	1.8	1.7	1.9	1.6	1.8
VDFS	Slope	S	2.6	2.5	2.7	2.9	2.4	2.9	2.1	2.9	2.5
VDFS	Slope	N	245	111	134	110	13[illegible]	58	53	52	82
VDFS	S-Intrcp	M	2.7	2.6	2.8	1.1	4.1	1.1	4.2	1.0	3.9
VDFS	S-Intrcp	S	4.5	4.3	4.6	3.8	4.5	3.8	4.3	3.9	4.6
VDFS	I-Intrcp	M	-0.6	-0.6	-0.6	0.2	-1.2	0.2	-1.4	0.2	-1.0
VDFS	I-Intrcp	S	2.2	2.2	2.2	1.7	2.4	1.6	2.4	1.7	2.4
LDCS	Slope	M	0.9	1.1	0.8	0.7	1.1	0.9	1.3	0.5	1.0
LDCS	Slope	S	1.3	1.5	1.2	1.3	1.4	1.5	1.4	1.0	1.3
LDCS	Slope	N	243	112	131	111	132	58	54	53	78
LDCS	S-Intrcp	M	0.1	-0.1	0.3	-0.3	0.4	-0.5	0.3	0.0	0.5
LDCS	S-Intrcp	S	2.4	2.4	2.4	1.3	3.0	1.3	3.2	1.3	2.9
LDCS	I-Intrcp	M	0.2	0.5	0.1	0.5	.0	0.7	0.2	0.3	-0.1
LDCS	I-Intrcp	S	2.1	1.9	2.2	1.3	2.6	1.0	2.6	1.5	2.5
LSPS	Slope	M	9.5	11.4	7.9	5.8	12.5	7.3	16.0	4.1	10.2
LSPS	Slope	S	14.2	14.8	13.5	12.5	14.9	13.5	14.9	11.0	14.5
LSPS	Slope	N	244	113	131	109	135	59	54	50	81
LSPS	S-Intrcp	M	4.0	1.0	6.6	-0.5	7.6	-2.5	4.8	1.9	9.5
LSPS	S-Intrcp	S	23.3	21.7	24.4	14.8	27.9	16.2	26.0	12.8	29.0
LSPS	I-Intrcp	M	0.1	0.3	0.0	0.2	0.1	0.5	0.1	-0.1	0.1
LSPS	I-Intrcp	S	1.8	1.8	1.8	1.3	2.1	1.2	2.2	1.2	2.0
SRDS	Slope	M	0.5	0.5	0.5	0.3	0.7	0.3	0.8	0.3	0.6
SRDS	Slope	S	0.6	0.6	0.6	0.4	0.7	0.4	0.7	0.4	0.6
SRDS	Slope	N	245	111	134	111	134	58	53	53	81
SRDS	S-Intrcp	M	-0.1	-0.2	0.0	-0.[illegible]	.0	-0.2	-0.1	-0.1	0.1
SRDS	S-Intrcp	S	1.2	1.1	1.3	0.5	1.6	0.5	1.5	0.5	1.7
SRDS	I-Intrcp	M	0.7	0.7	0.6	0.6	0.8	0.6	0.8	0.5	0.7
SRDS	I-Intrcp	S	1.5	1.4	1.5	1.3	1.6	1.5	1.4	1.1	1.7

1 3

Table 33

Interactive Reading Assessment System - Spanish:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Comprehension Scales Overall and by Language Category

Scale	Measure	Statistic	Overall	Low Eng	High Eng	Low Span	High Span	Lo Lo	Lo Hi	Hi Lo	Hi Hi
NRCS	Slope	M	0.9	0.9	0.9	0.3	1.4	0.3	1.6	0.4	1.2
NRCS	Slope	S	1.4	1.4	1.3	0.9	1.5	0.9	1.6	0.9	1.4
NRCS	Slope	N	248	113	135	112	136	59	54	53	82
NRCS	S-Intrcp	M	-1.0	-1.1	-1.0	-0.6	-1.4	-0.6	-1.7	-0.5	-1.3
NRCS	S-Intrcp	S	2.6	2.8	2.5	2.0	3.0	2.1	3.3	1.8	2.8
NRCS	I-Intrcp	M	0.7	0.8	0.6	0.3	1.0	0.3	1.2	0.4	0.8
NRCS	I-Intrcp	S	1.2	1.2	1.2	1.0	1.3	1.0	1.2	0.9	1.3
ERCS	Slope	M	0.9	0.8	1.0	0.3	1.4	0.2	1.6	0.4	1.3
ERCS	Slope	S	1.5	1.5	1.5	0.9	1.6	0.8	1.7	0.9	1.6
ERCS	Slope	N	245	111	134	111	134	58	53	53	81
ERCS	S-Intrcp	M	-1.2	-1.1	-1.4	-0.5	-1.8	-0.4	-1.8	-0.5	-1.9
ERCS	S-Intrcp	S	2.9	2.8	2.9	1.7	3.4	1.9	3.4	1.5	3.4
ERCS	I-Intrcp	M	0.7	0.6	0.7	0.3	1.0	0.3	1.0	0.4	1.0
ERCS	I-Intrcp	S	1.4	1.3	1.5	0.9	1.7	0.9	1.5	1.0	1.8
NLCS	Slope	M	0.8	0.9	0.8	0.6	1.1	0.7	1.2	0.4	1.0
NLCS	Slope	S	1.2	1.2	1.1	1.2	1.1	1.1	1.2	1.3	2.0
NLCS	Slope	N	246	111	135	112	134	59	52	53	82
NLCS	S-Intrcp	M	1.7	1.8	1.7	0.9	2.4	1.0	2.6	0.7	2.3
NLCS	S-Intrcp	S	2.3	2.3	2.2	2.0	2.2	2.0	2.3	1.9	2.2
NLCS	I-Intrcp	M	-0.8	-0.9	-0.8	-0.1	-1.4	-0.2	-1.7	.0	-1.3
NLCS	I-Intrcp	S	2.2	2.3	2.2	1.9	2.3	2.0	2.3	1.7	2.3
ELCS	Slope	M	0.9	0.9	0.9	0.5	1.2	0.5	1.3	0.5	1.1
ELCS	Slope	S	1.5	1.5	1.5	1.4	1.5	1.3	1.6	1.4	1.4
ELCS	Slope	N	245	111	134	111	134	58	53	53	81
ELCS	S-Intrcp	M	0.7	0.7	0.7	0.1	1.3	0.2	1.3	-0.1	1.2
ELCS	S-Intrcp	S	2.8	2.8	2.8	2.2	3.0	2.5	2.9	1.9	3.1
ELCS	I-Intrcp	M	-0.1	-0.1	-0.1	0.3	-0.5	0.1	-0.4	0.6	-0.5
ELCS	I-Intrcp	S	2.1	2.1	2.1	1.6	2.4	2.0	2.3	0.9	2.5

level below expectation for both comprehension scales. Thus, under the growth track model, these bilingual students, in the aggregate, enter first grade with Spanish oral skills which exceed the expectations of the growth track model, as was found for their English skills. Their decoding skills, also as in English, have yet to be developed.

The instruction-intercepts reveal a similar pattern: negative values for the formal language tasks, and positive values, between 0 and 1, for those tasks requiring decoding skills. Again, due to the handling of special cases for the instruction-intercepts, these values will not exactly correspond to those that would be expected given the slope and student-intercept indices.

For the slope indices for the formal language tasks, growth is slightly below expectation in Vocabulary Definitions, and about half the level expected for the listening comprehension tasks. For the decoding tasks, growth is slightly above expectations in Real Word Decoding -- for the other two decoding tasks, were the growth track model is not strictly applicable, growth is at one level per year for Synthetic Word Decoding, and 10% per year for Synthetic Word Decoding. For Sentence Reading, growth proceeds at about a half syllable per year, and at only half the level expected for the two comprehension scales.

The aggregate picture is one where the acquisition of reading comprehension is well below the expectations of the growth track model. Decoding skills for isolated words are close to the growth track, though this skill is not as great as evidenced in the English materials. Oral language skills, which are above expectation at entry to first grade, show growth that is substantially below expectation -- as such, these formal language skills are projected to fall below the growth track, and thus, would seem to provide a major obstacle to Spanish reading acquisition.

Spanish analyses of variance. As was done for the English scales, to assess whether or not the overall description differs for the various entry language groups and sites, a series of analyses of variance were conducted on the scale growth measures: each of the three main summary measures of growth for each of the nine IRAS scales were subjected to a 2 x 2 x 2 x 5 between-subject analysis, comprising the variables of English entry (low and high), Spanish entry (low and high), gender (boy and girl), and site (0, 1, 2, 3, and 5). Again, only those main effects significant at the .05 level will be discussed, and only those two-way interactions significant at the .01 level.

The results of these analyses are summarized in Table 34 in a format matching that used in presenting the English IRAS analyses of variance summaries. Given that not all of the variables showed systematic effects, only a subset have been tabled: the main effects of Spanish entry and site, and the interactions of English entry by Spanish entry and Spanish entry by site. The tables are identically organized to those presenting the English summaries. Extracts from the

Table 34

Interactive Reading Assessment System - Spanish:
Summary of the Analyses of Variance
on Growth Indices for the Bilingual Sample

Scale - Measure		Variables: F-ratio/p-value				Residual Sum Sqrs	Sum Sqrs Explain/ Residual
		Spanish	Site	Eng x Spn	Spn x Site		
VDC	Slp	2.3/NS	5.1/.001	.6/NS	1.8/NS	1528.8	.26
	IIn	6.1/.014	4.2/.003	4.9/NS	1.0/NS	3569.5	.18
	XIn	4.6/.033	3.6/.008	1.1/NS	.3/NS	561.3	.14
VDF	Slp	2.0/NS	3.8/.005	.2/NS	1.2/NS	1461.0	.13
	IIn	18.2/.001	6.3/.001	2.9/NS	.6/NS	3635.9	.33
	XIn	9.6/.002	5.7/.001	5.5/NS	.6/NS	922.3	.30
LDC	Slp	1.3/NS	11.3/.001	.2/NS	5.4/.001	307.4	.41
	IIn	10.8/.001	5.8/.001	3.5/NS	3.5/.008	1109.6	.28
	XIn	8.9/.003	5.6/.001	5.2/NS	2.5/NS	808.3	.30
LSF	Slp	1.2/NS	8.0/.001	1.6/NS	1.8/NS	36919.0	.33
	IIn	5.6/.019	1.3/NS	1.3/NS	.8/NS	117214.2	.12
	XIn	1.6/NS	1.5/NS	6.5/NS	1.0/NS	677.5	.11
SRD	Slp	3.8/NS	6.9/.001	2.5/NS	.7/NS	62.9	.37
	IIn	2.1/NS	3.5/.009	.4/NS	.4/NS	316.5	.16
	XIn	2.7/NS	5.8/.001	.1/NS	1.7/NS	434.5	.22
NRC	Slp	13.0/.001	13.2/.001	3.3/NS	1.4/NS	286.7	.62
	IIn	.3/NS	17.7/.001	.5/NS	2.4/NS	1137.5	.50
	XIn	2.7/NS	26.9/.001	.4/NS	1.6/NS	204.7	.75
ERC	Slp	13.5/.001	8.7/.001	8.7/.003	1.5/NS	327.7	.59
	IIn	1.4/NS	11.4/.001	.3/NS	1.2/NS	1383.8	.44
	XIn	1.2/NS	14.7/.001	2.0/NS	1.1/NS	323.3	.50
NLC	Slp	.2/NS	6.2/.001	.5/NS	1.2/NS	262.0	.23
	IIn	13.9/.001	3.1/NS	1.1/NS	2.3/NS	937.1	.33
	XIn	13.7/.001	1.9/NS	.4/NS	.8/NS	933.1	.29
ELC	Slp	.4/NS	5.1/.001	.3/NS	.2/NS	422.3	.23
	IIn	9.5/.002	1.4/NS	.0/NS	.4/NS	1566.5	.19
	XIn	9.7/.002	1.8/NS	.2/NS	1.6/NS	852.1	.26

Note: Significance levels are .05 for main effects, .01 for interactions.

full set of ANOVA summary tables used in constructing Table 34 are presented in Appendix E.

The summary tables reveal that the most significant variables from the set for the Spanish IRAS scales are those of Spanish entry and site. Thus, the discussion which follows will concentrate on the differences expressed in these variables, treating Spanish entry first.

Spanish entry is found to have significant effects on the slope indices for the reading comprehension scales, and on the intercepts for Real Word Decoding, Vocabulary Definition, both synthetic word tasks, and the two listening comprehension tasks. Figures 31 through 39 display the growth functions for the two Spanish entry groups, along with the overall growth functions, for each of the nine IRAS scales (the data are taken from Tables 32 and 33).

For the formal language tasks, significant differences between the Spanish entry groups are found for the intercept indices, but not for the slopes. For Vocabulary Definitions, displayed in Figure 32, the low Spanish group closely follows the growth track prediction, with the high Spanish group about three IRAS levels above it. For Narrative Listening Comprehension (Figure 38), both groups show growth rates of about one IRAS level per year (half that expected), with the high Spanish group about one and a half levels above the low Spanish group at entry to first grade. A similar pattern holds for Expository Listening Comprehension (Figure 39), but the entry points are about one IRAS level lower. Overall, the high Spanish group shows a substantial advantage over the low Spanish group at first grade entry, and the growth rates, which do not differ, allow the advantage to be maintained. However, growth in text comprehension is at level substantially below expectation (and below that expected from the sample's ability shown in word knowledge), and thus, the 'text' skills of the sample are not projected to be sufficient to support sustained Spanish reading growth for either entry group -- the low Spanish group is predicted to fall below the growth track around the middle of first grade, while the high Spanish group does not do so until the middle of third grade.

For decoding skills, as in the formal language skills, Spanish entry skill has significant effects on the intercept values, but not the slopes. For Real Word Decoding (Figure 31), the low Spanish group enters first grade at about a quarter of an IRAS level below the growth track prediction of zero, and the high Spanish group, about three quarters of a level above. Thus, given similar growth rates, the low Spanish group is predicted to remain below the growth track while the high Spanish group above it. For Synthetic Word Decoding (Figure 33), the high Spanish group reaches the limits of the test by fourth-grade exit. The student-intercept measures for the two groups are close to those found in Real Word Decoding, with a growth rate of about one IRAS level per year. For Synthetic Word Spelling, performance is similar to that found in English with both groups well below the limits of the test by fourth-grade exit. The low Spanish group enters first grade with no skill in this area, while the high Spanish group enters

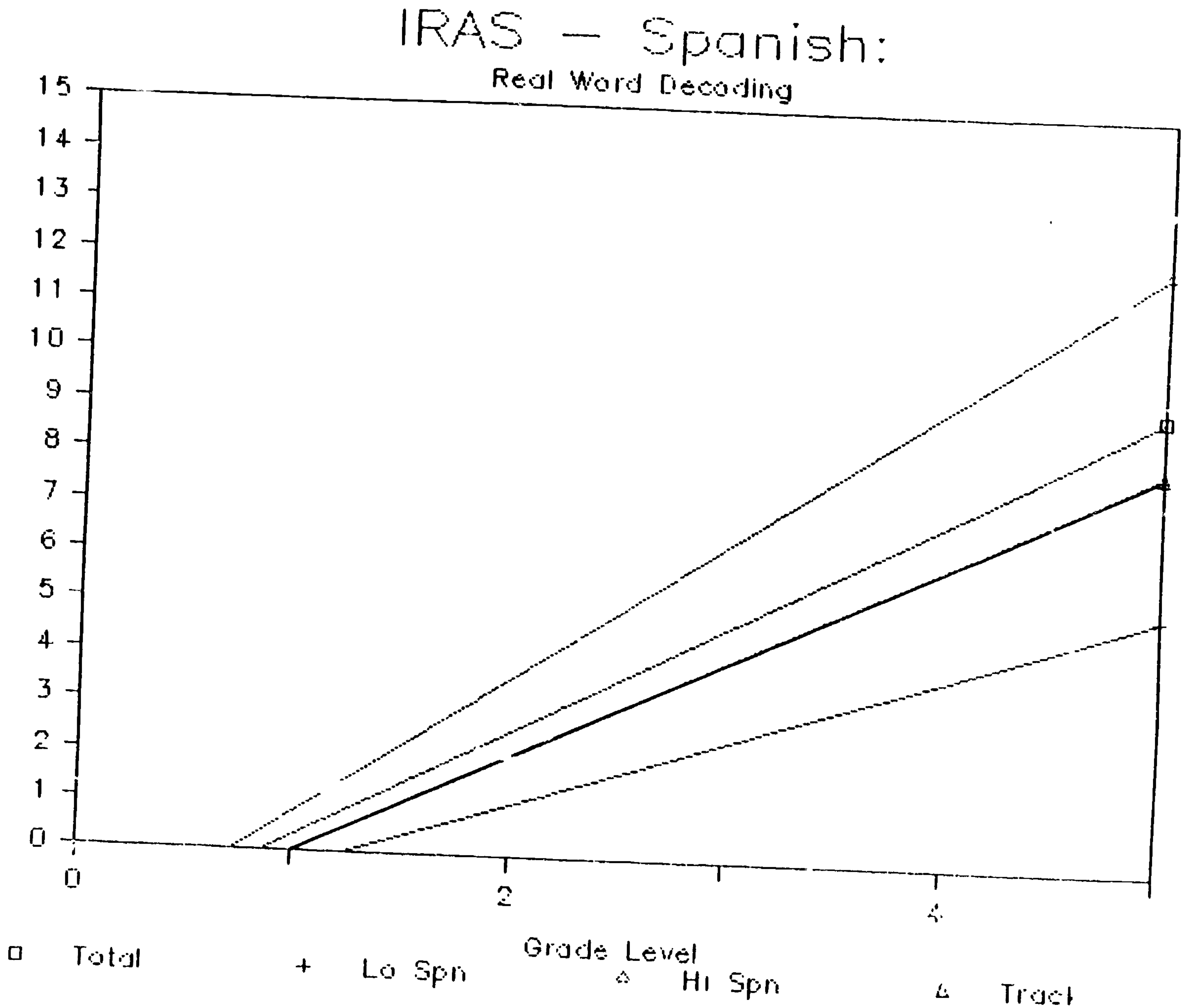


Figure 31. IRAS-S Real Word Decoding growth for the entire bilingual sample and by Spanish entry category.

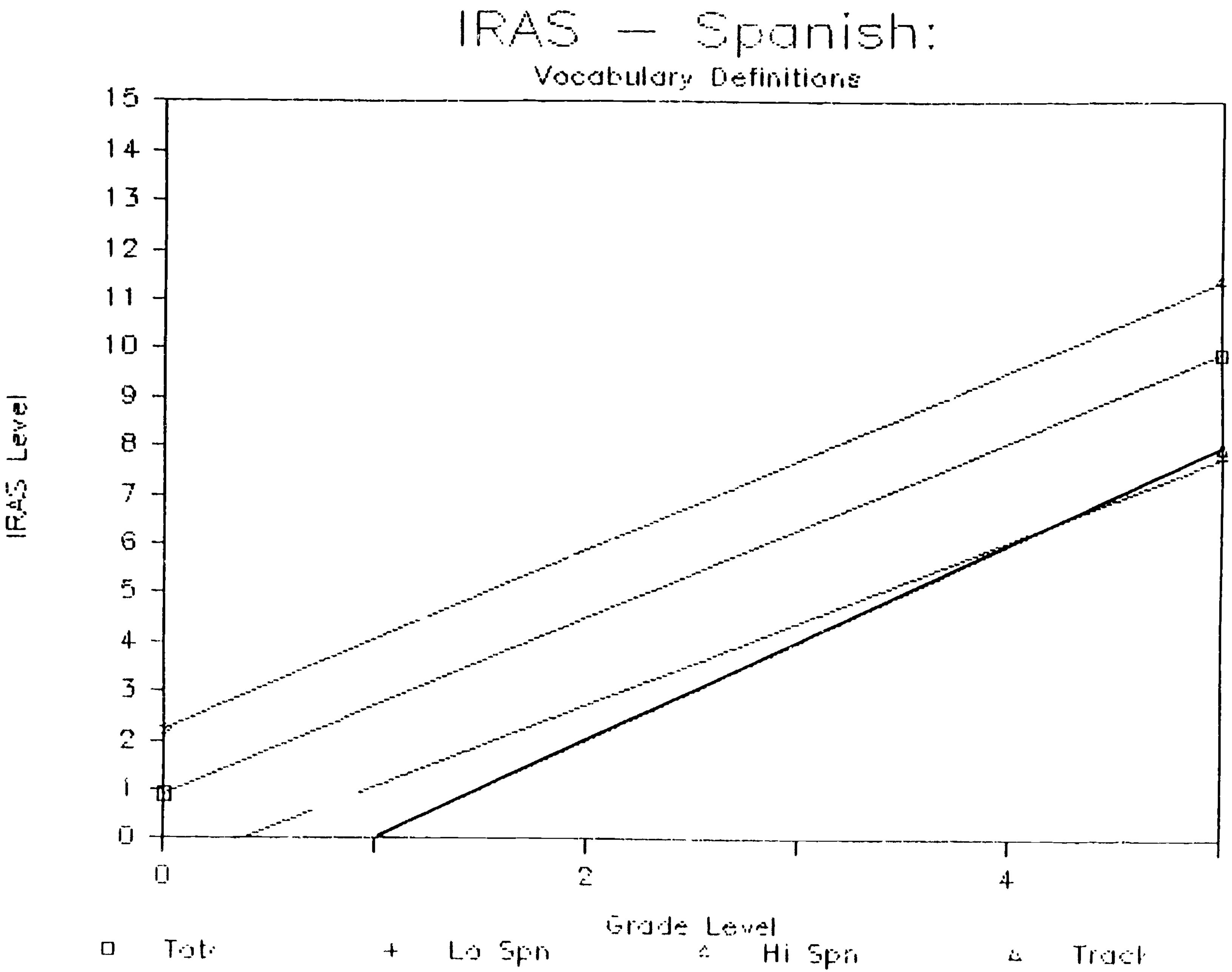


Figure 32. IRAS-S Vocabulary Definition growth for the entire bilingual sample and by Spanish entry category.

110

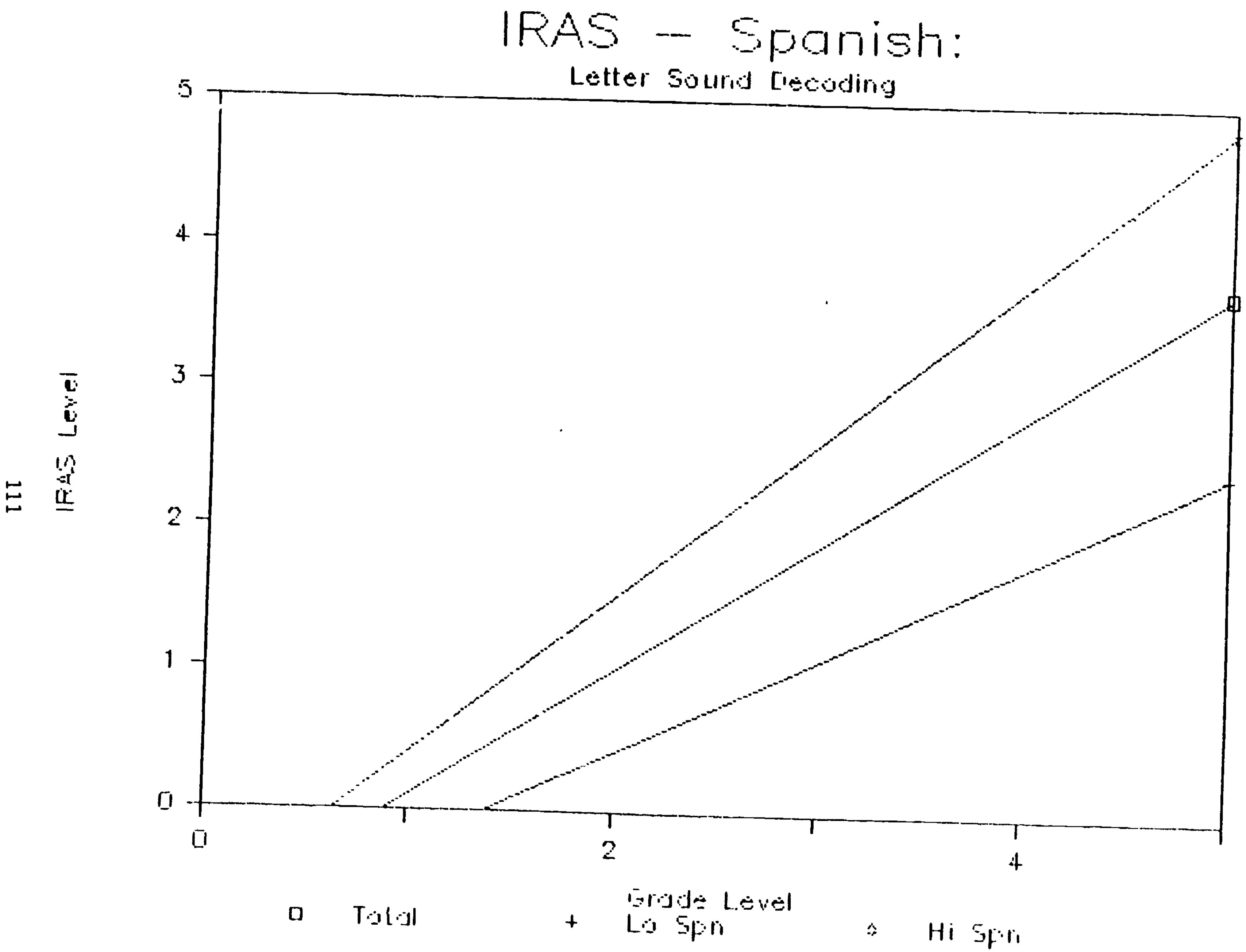


Figure 33. IRAS-S Synthetic Word Decoding growth for the entire bilingual sample and by Spanish entry category.

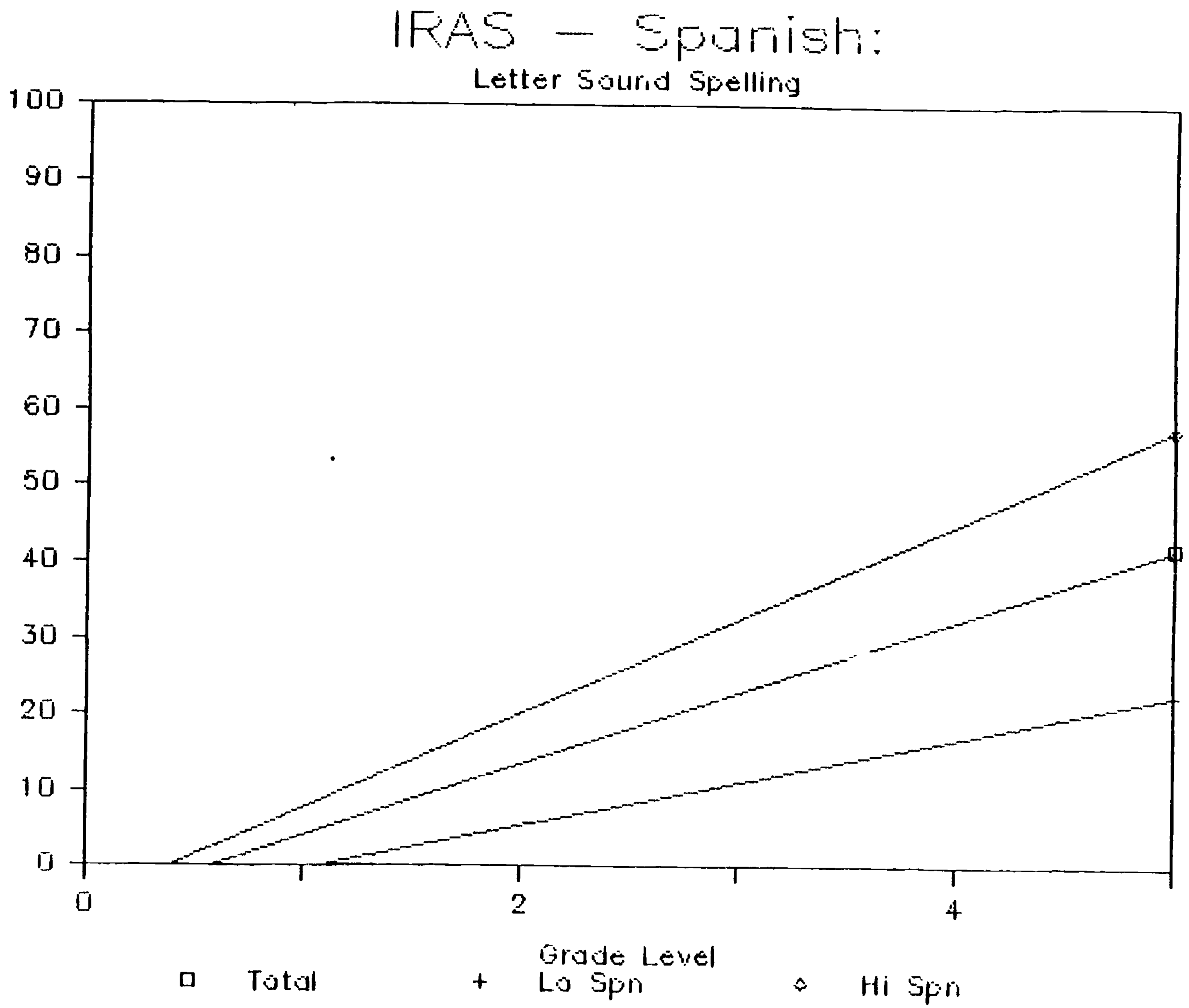
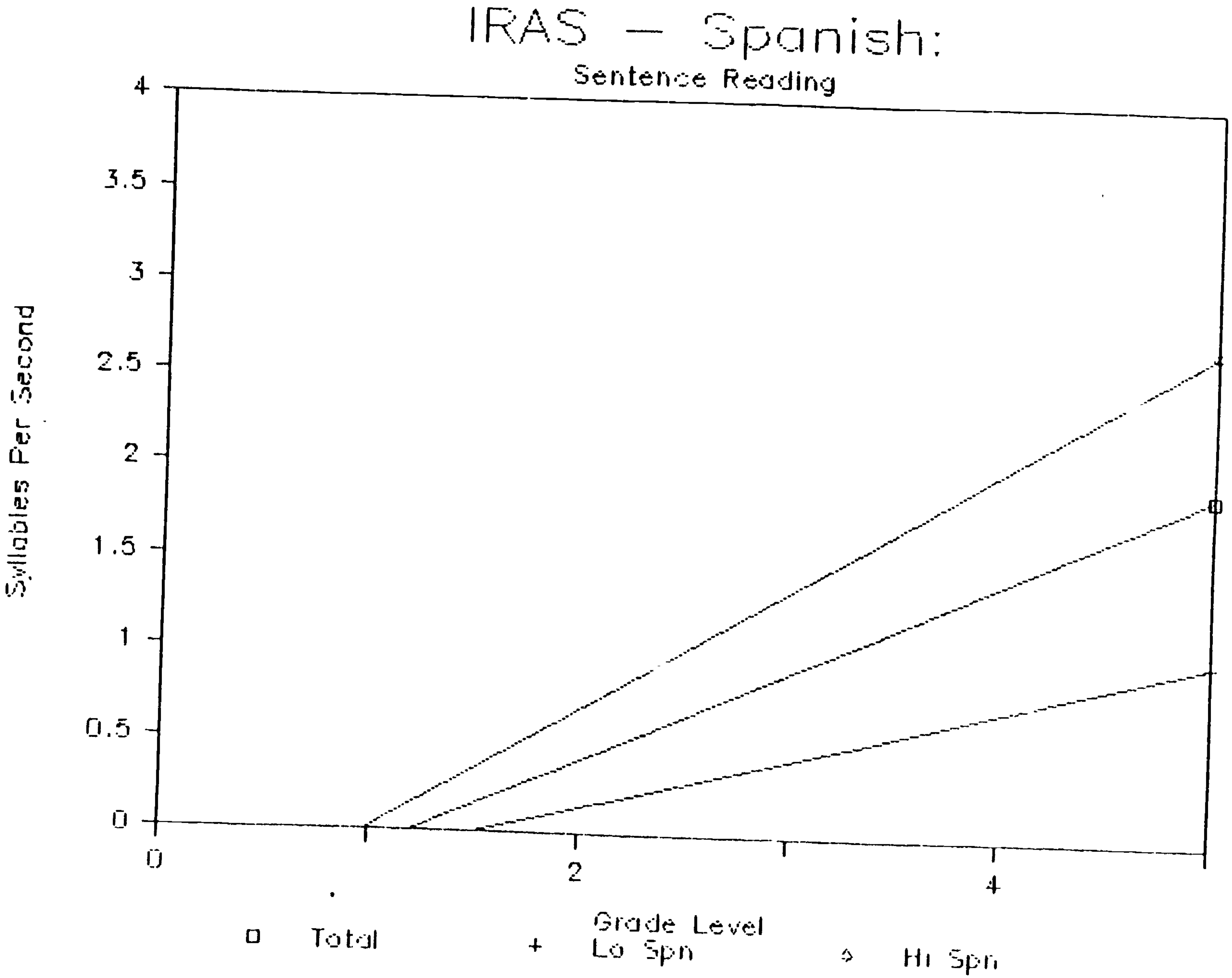


Figure 34. IRAS-S Synthetic Word Spelling growth for the entire bilingual sample and by Spanish entry category.



113

Figure 35. IRAS-S Sentence Reading growth for the entire bilingual sample and by Spanish entry category.

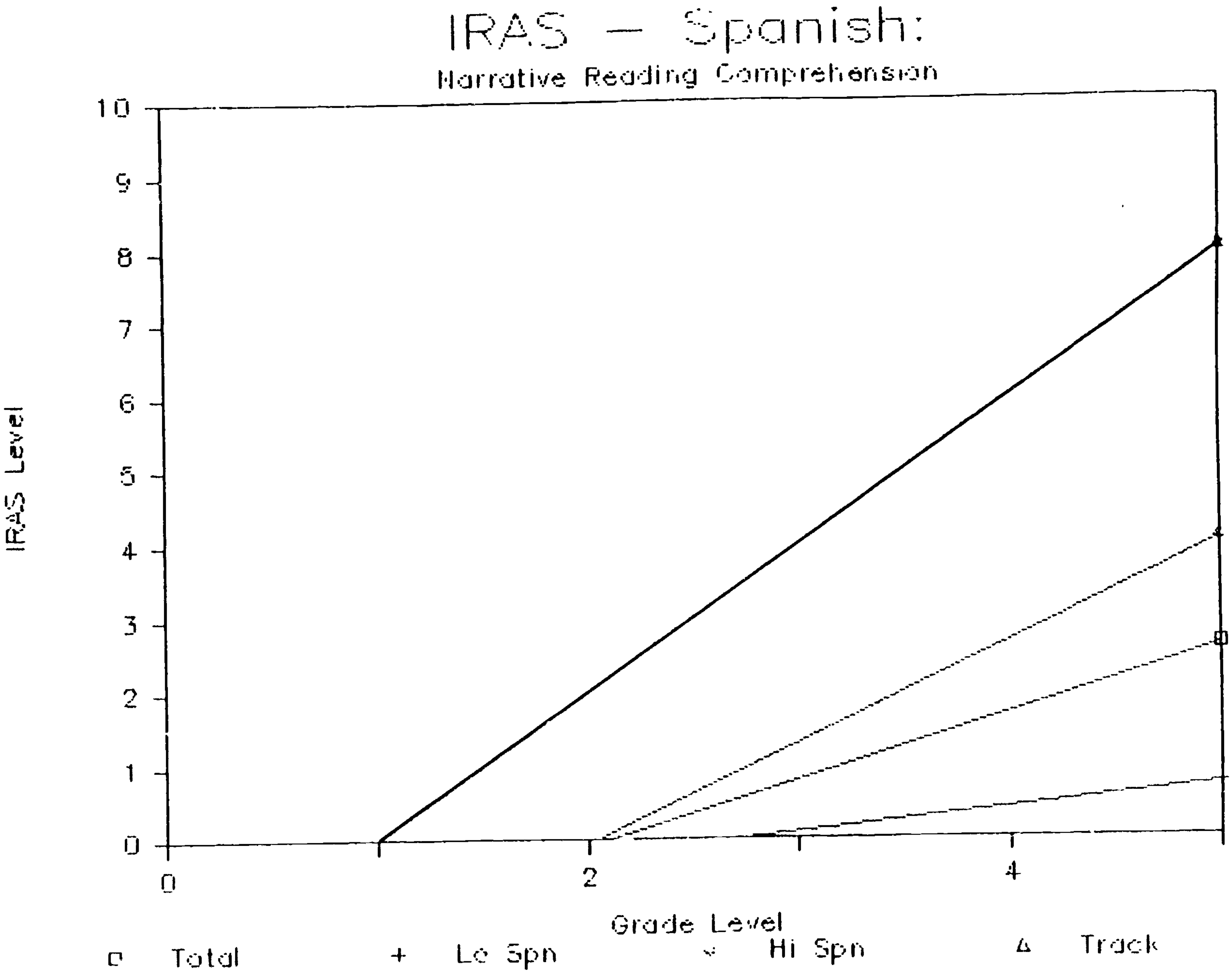


Figure 36. IRAS-S Narrative Reading Comprehension growth for the entire bilingual sample and by Spanish entry category.

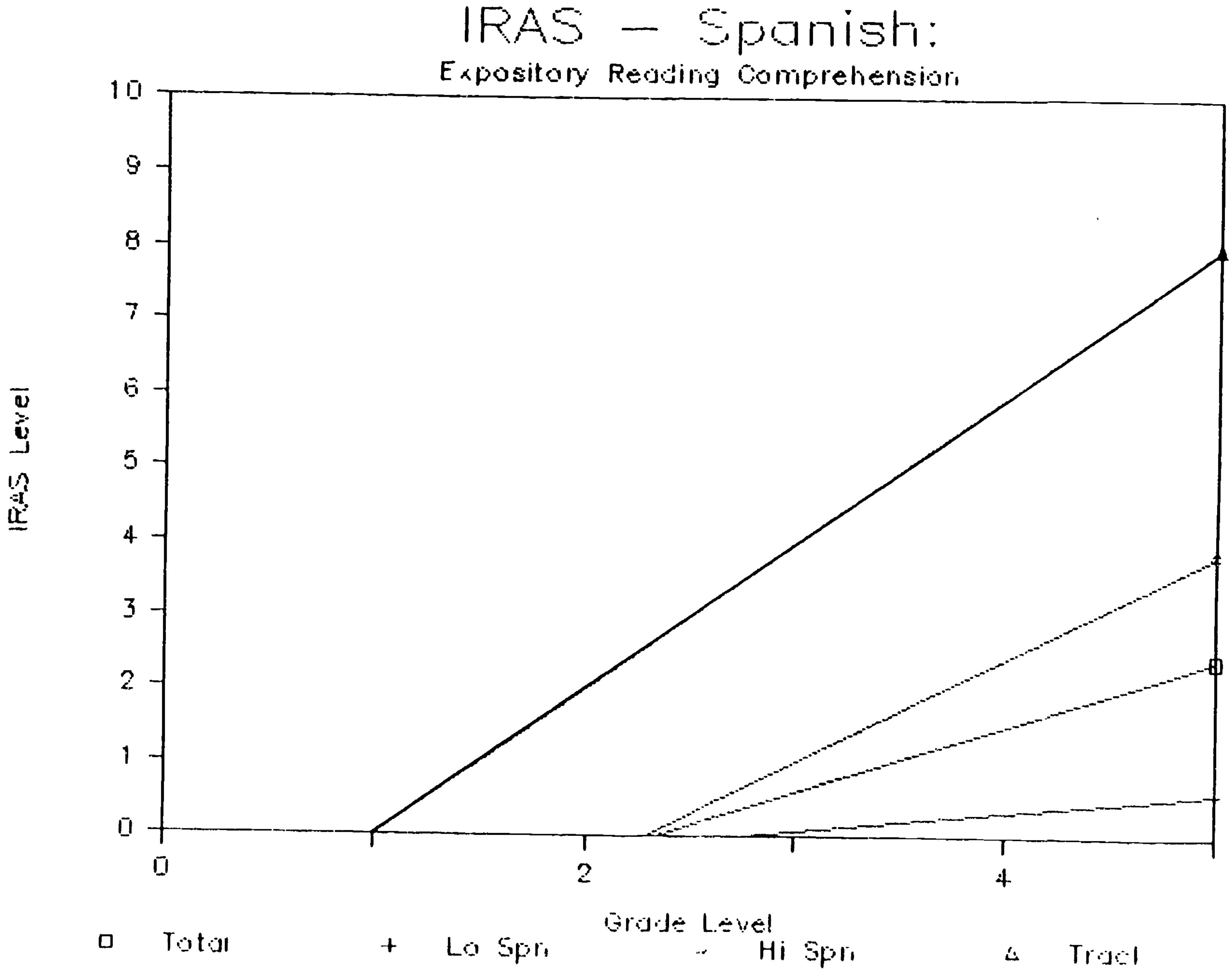


Figure 37. IRAS-S Expository Reading Comprehension growth for the entire bilingual sample and by Spanish entry category.

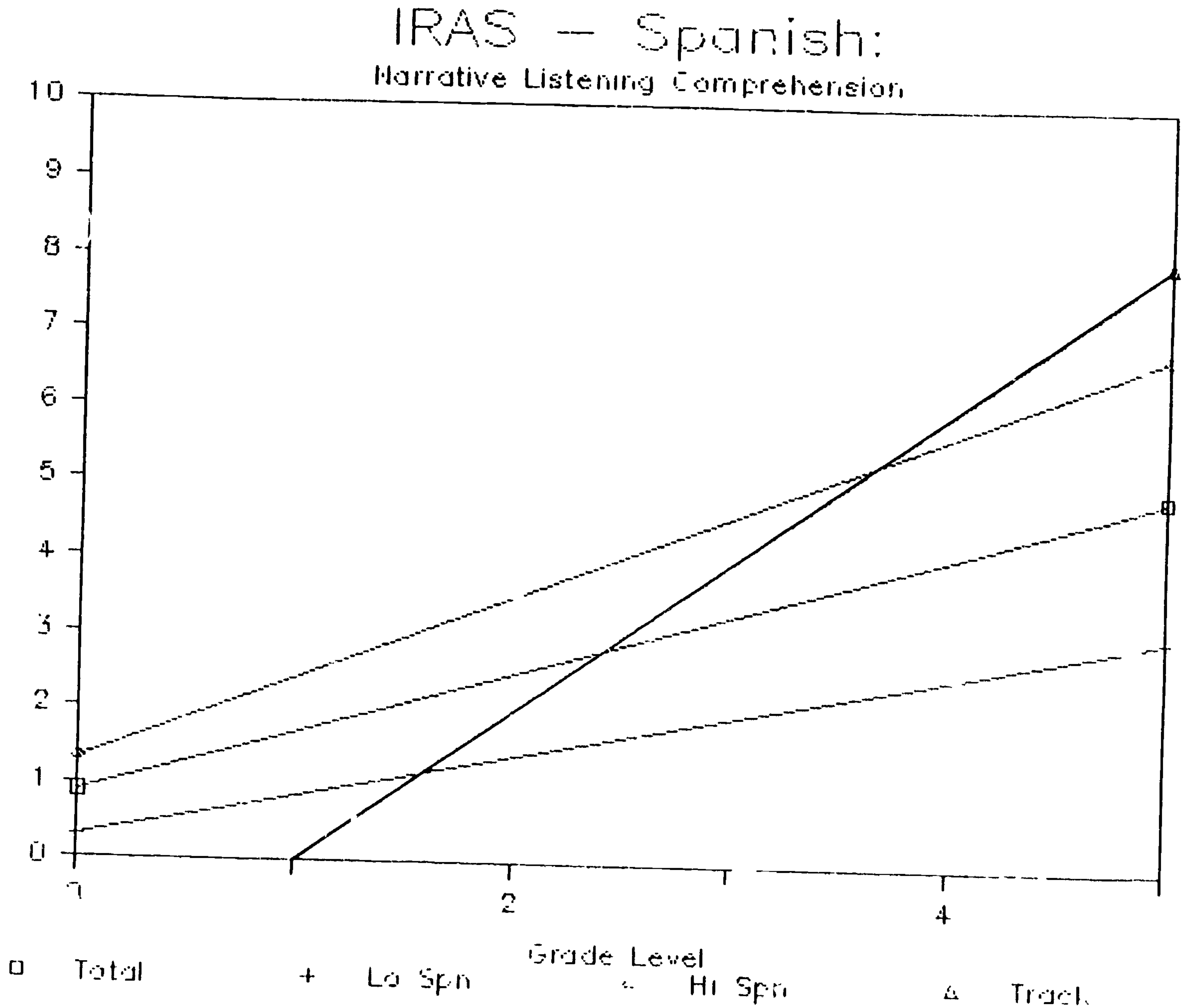


Figure 38. IRAS-S Narrative Listening Comprehension growth for the entire bilingual sample and by Spanish entry category.

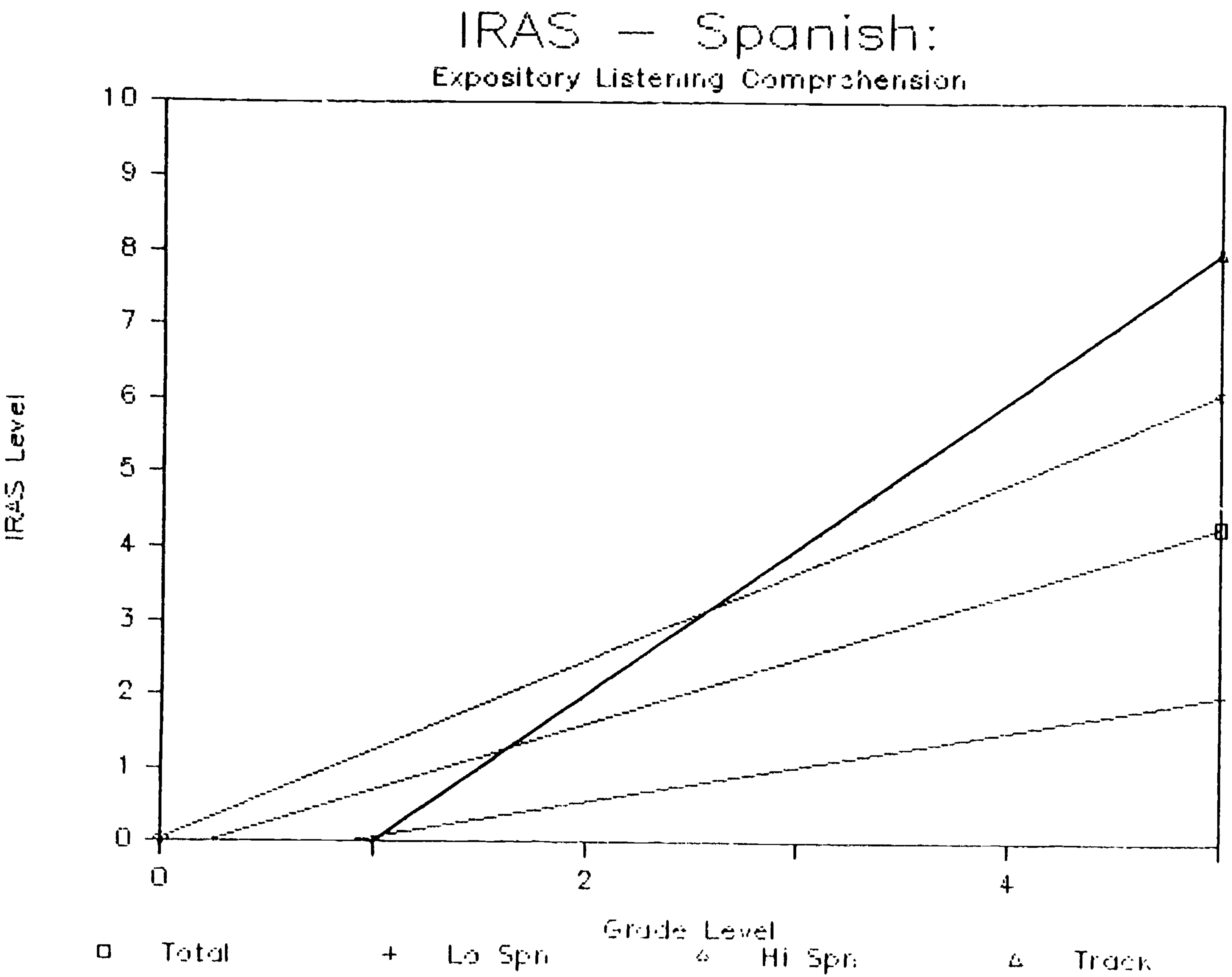


Figure 39. IRAS-S Expository Listening Comprehension growth for the entire bilingual sample and by Spanish entry category.

with the ability to correctly spell about 10% of the words. Thus, decoding of isolated items is difficult for the low Spanish entry group, and would be expected to be a barrier to progress in Spanish reading acquisition. Such is not the case for the high Spanish group, however.

For the task of Sentence Reading (Figure 35), none of the growth indices are influenced by Spanish entry skill -- the average growth rate of .5 syllables per second, starting from a near zero entry point, predicts a relatively low fluency rate of about 2 syllables per second on fourth-grade exit. For Narrative Reading Comprehension (Figure 36), only the slope indices are influenced by Spanish entry skill. Both groups show student-intercepts about a level below the growth track prediction, with growth for the low Spanish entry group at about a third of an IRAS level per year, while that of the high Spanish group is a full level above this rate. This same pattern holds for Expository Reading Comprehension (Figure 37). Thus, for the reading tasks, for the low Spanish group, little progress is made in acquiring Spanish reading skill; for the high Spanish group, progress is greater, but substantially below the growth track predictions.

Overall, the differences due to Spanish entry suggest that the low Spanish group enters first grade with less skill than the high Spanish group in the areas of formal language and decoding, but that subsequent growth does not differ. (Alternatively, looking at the instruction-intercepts, effective instruction seems to begin earlier for the high Spanish entry group than for the low Spanish group.) For reading comprehension, however, the two groups begin with the same low-level skills (or receive instruction at the same point in time), but, given the greater formal language and decoding skills of the high Spanish entry group their growth in reading comprehension is able to proceed at a greater rate. This rate, however, is substantially below that expected from the growth track model, and the data suggest that the major difficulty for these students is not decoding skill (though fluency may present some difficulties), but rather, skill in dealing with the formal language aspects of text.

As seen in Table 34, the variable of site, as in the English data, is also found to have widespread effects on Spanish performance, covering all of the IRAS scales and all of the growth measures (except the intercept indices in Synthetic Word Spelling, and Narrative and Expository Listening Comprehension). Figures 40 through 48 display the growth functions for the five sites for each of the nine scales (the data are taken from Appendix D). In describing the site differences each of the scales will be discussed individually, treating those dealing with formal language, decoding, and reading in turn.

For the formal language task of Vocabulary Definition (Figure 41), Sites 0 and 1 show high student-intercepts that are about two grade levels (4 IRAS levels) above expectation. Sites 2 and 3 show the lowest student-intercepts, at about one IRAS level, with Site 5 midway between these extremes. Growth for the border sites (0, 1, and 2) is slightly below the expected rate of two levels per year (1.7), while

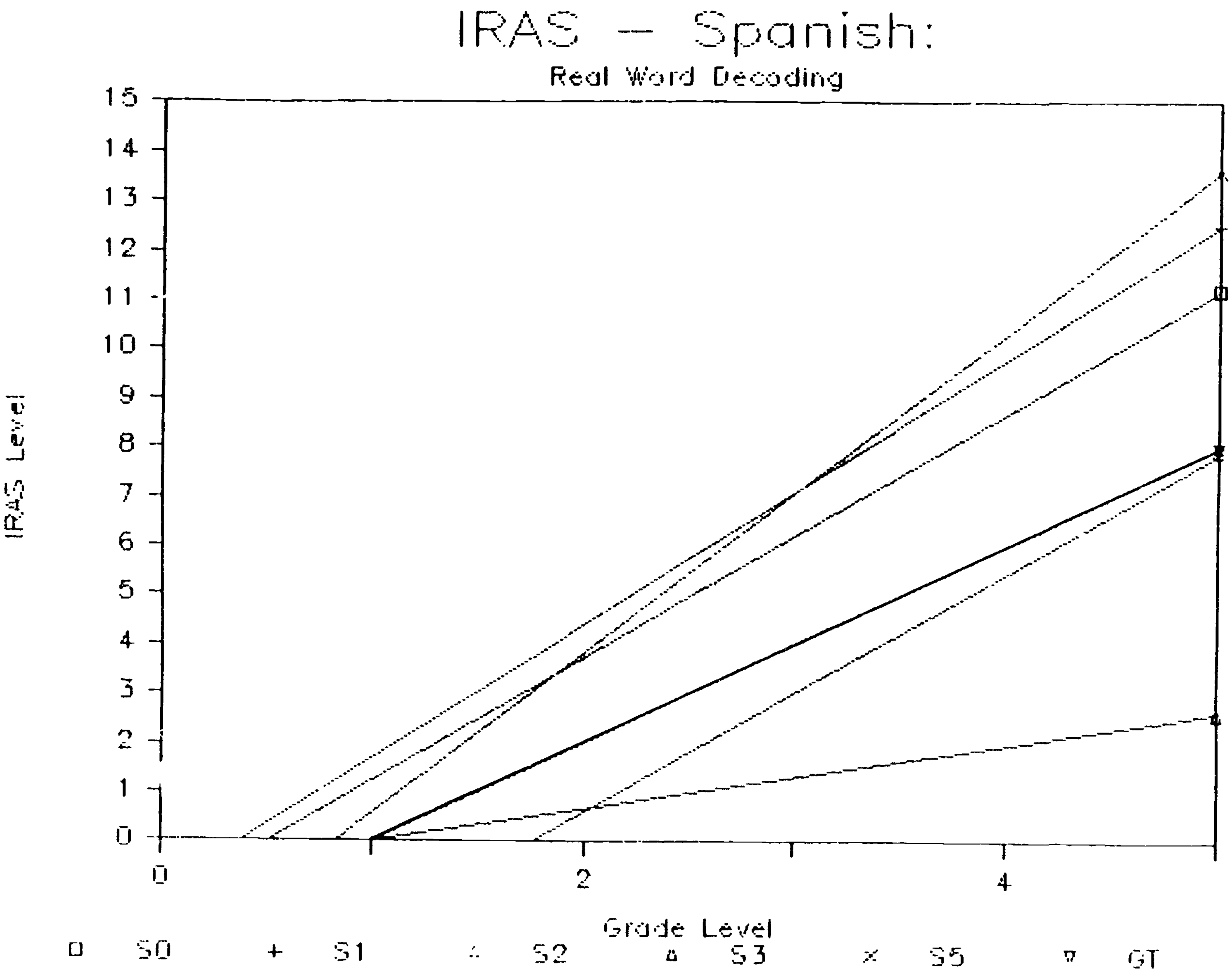


Figure 40. IRAS-S Real Word Decoding growth for each site.

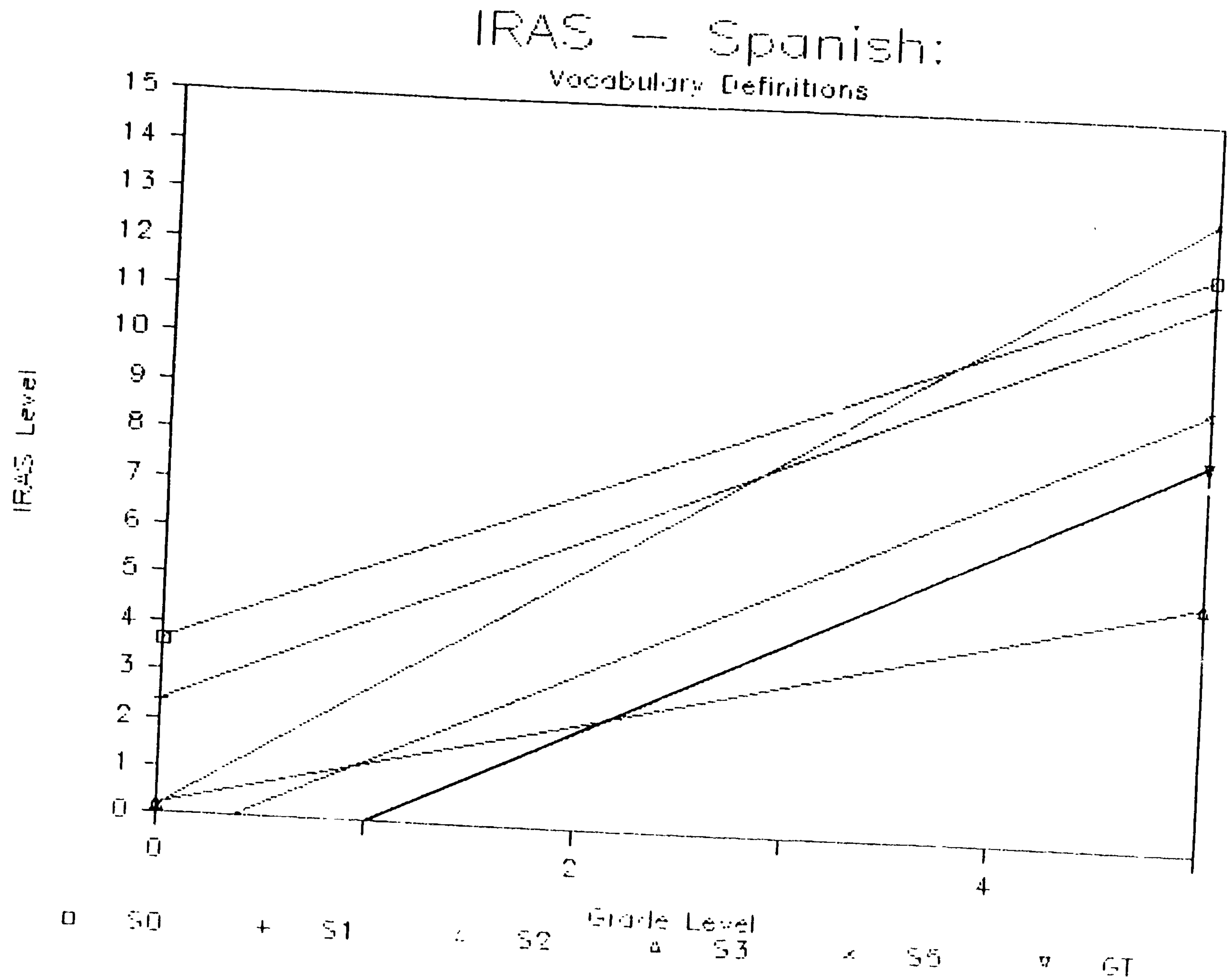


Figure 41. IRAS-S Vocabulary Definition growth for each site.

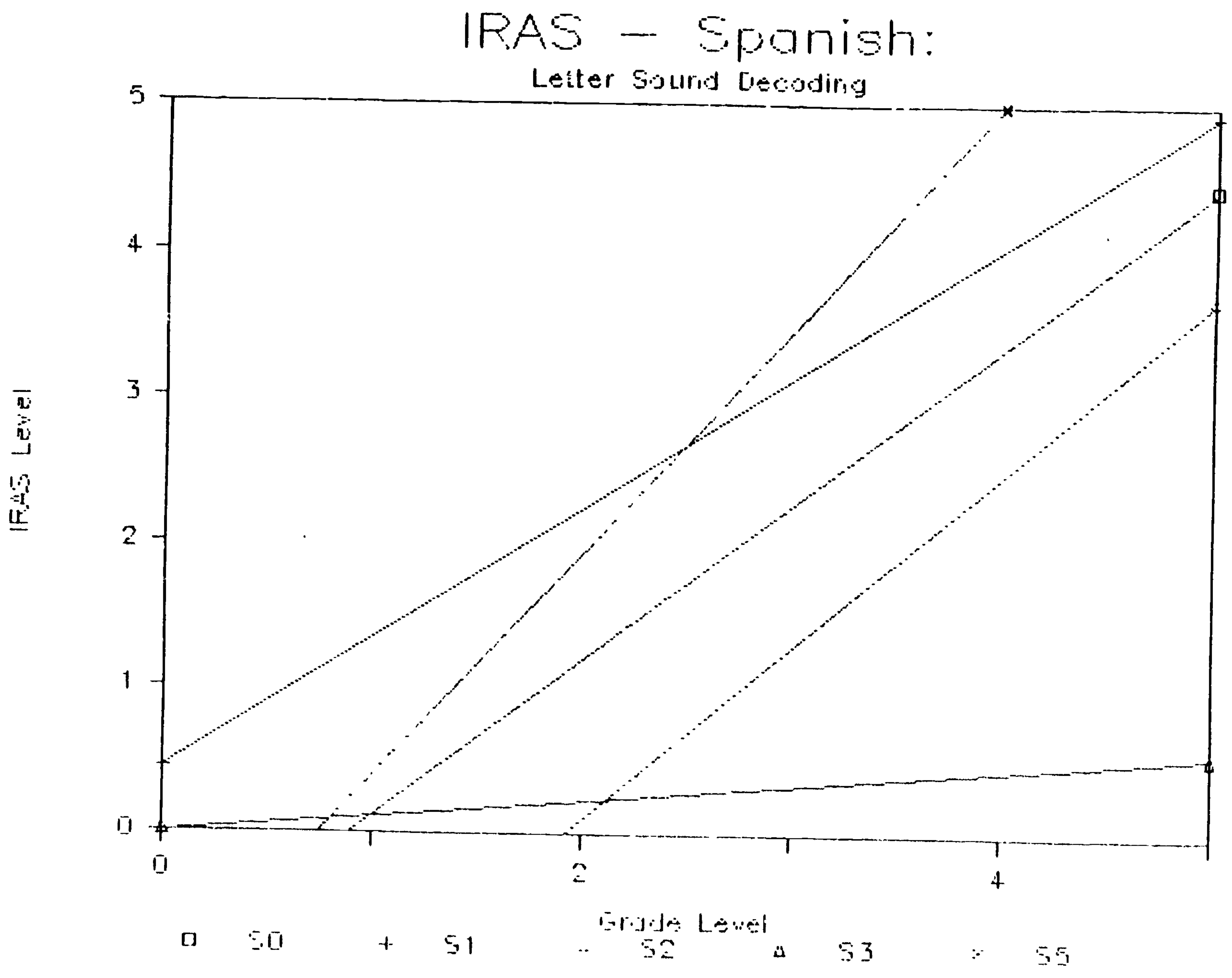


Figure 42. IRAS-S Synthetic Word Decoding growth for each site.

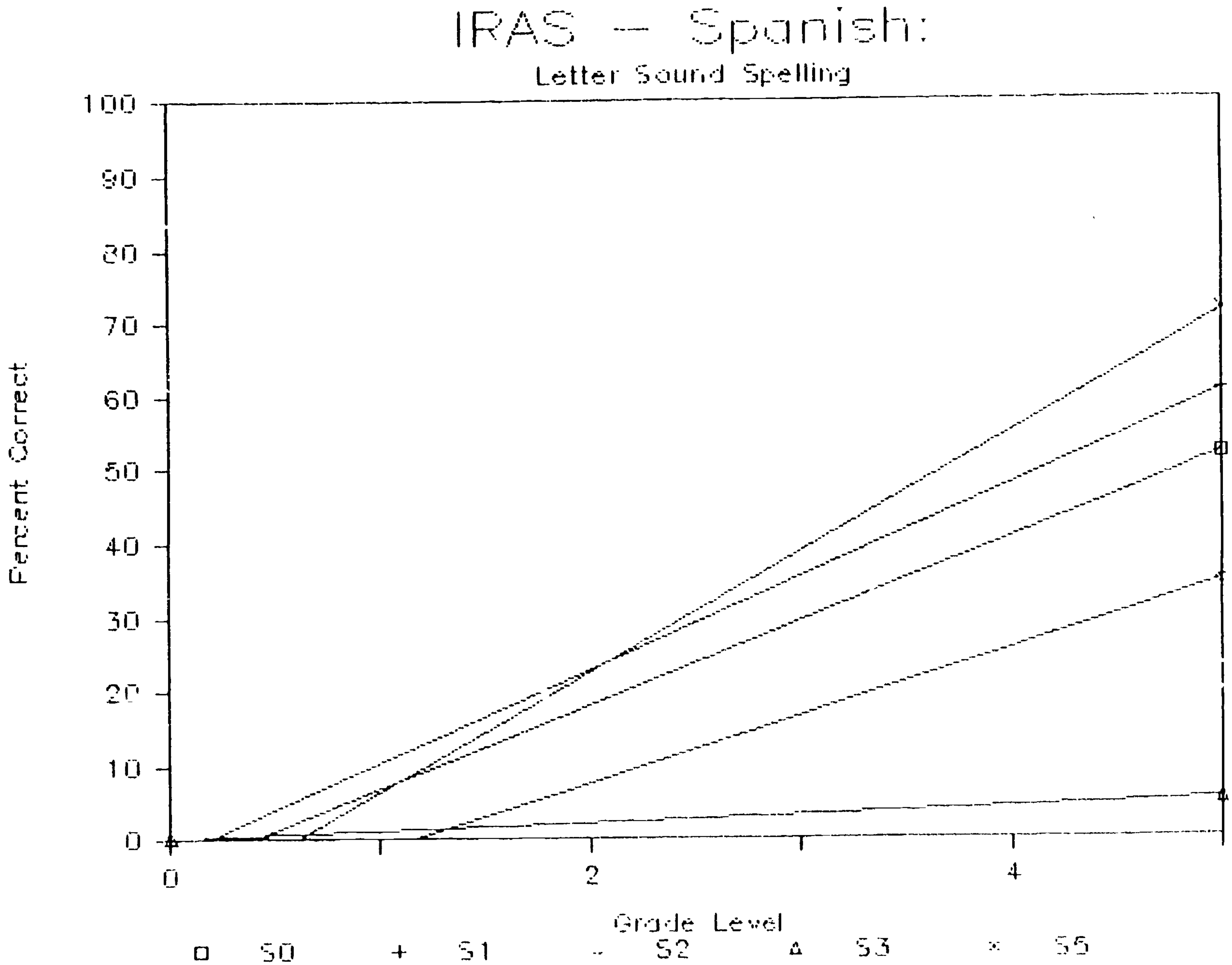


Figure 43. IRAS-S Synthetic Word Spelling growth for each site.

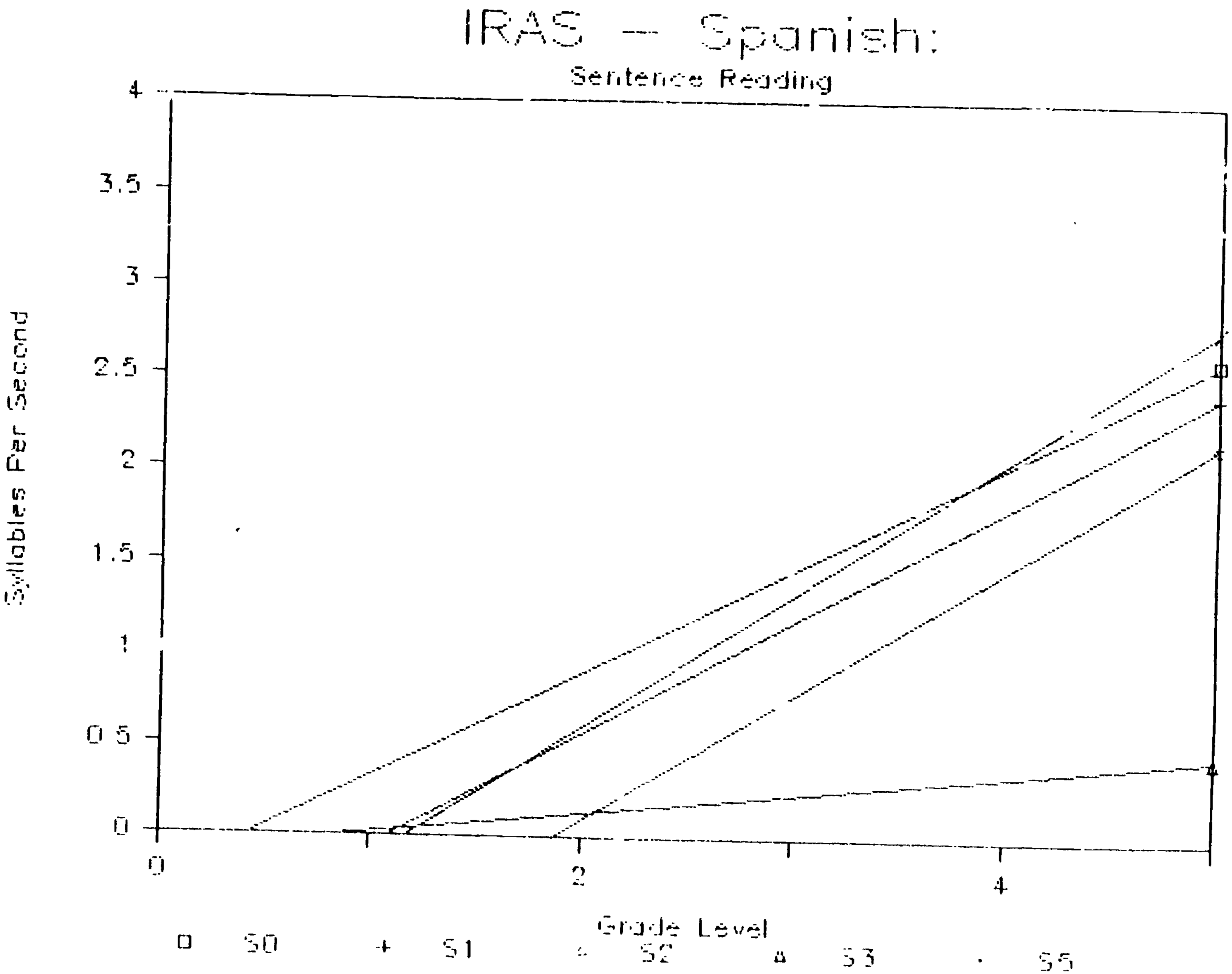


Figure 44. IRAS-S Sentence Reading growth for each site.

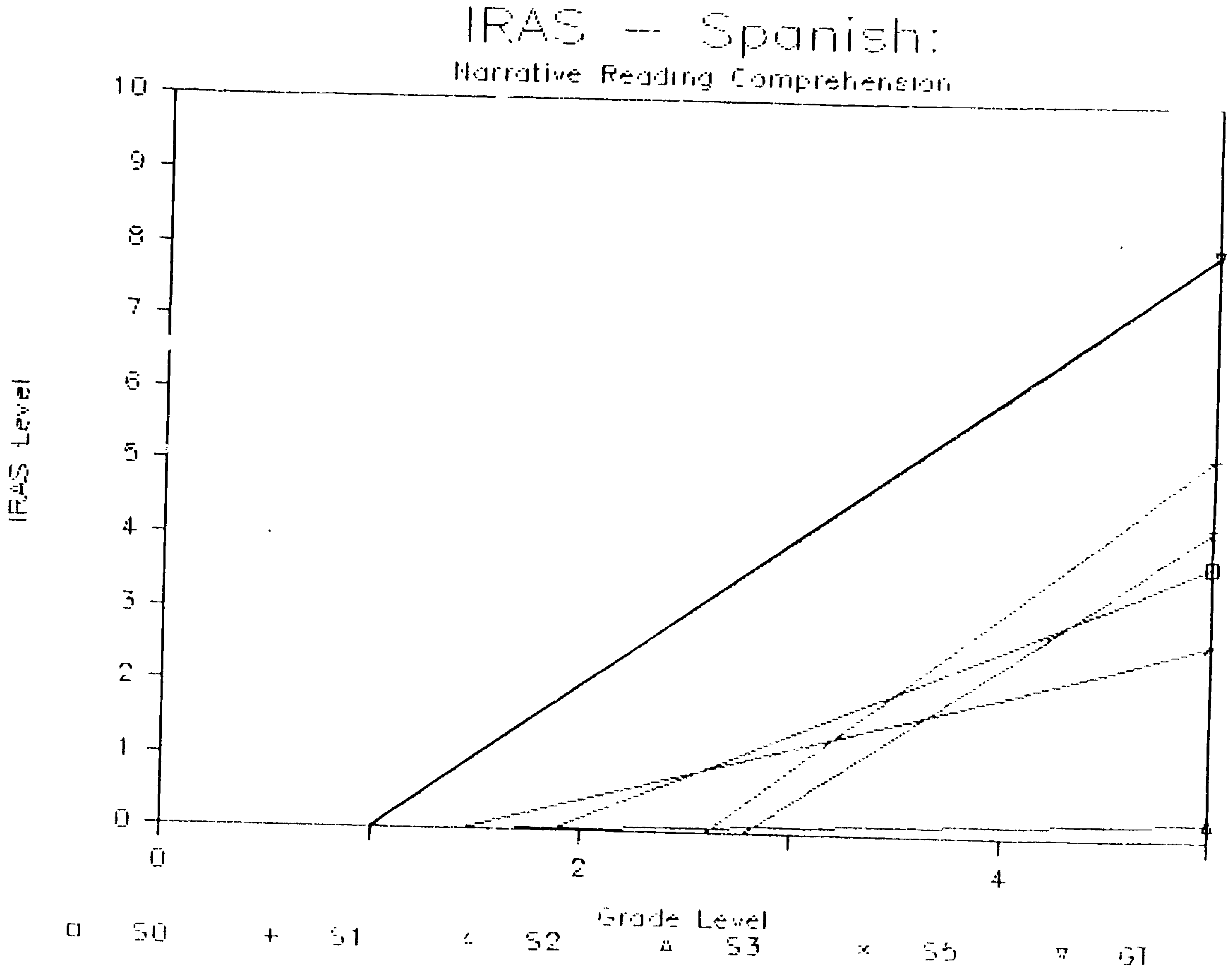


Figure 45. IRAS-S Narrative Reading Comprehension growth for ach site.

125

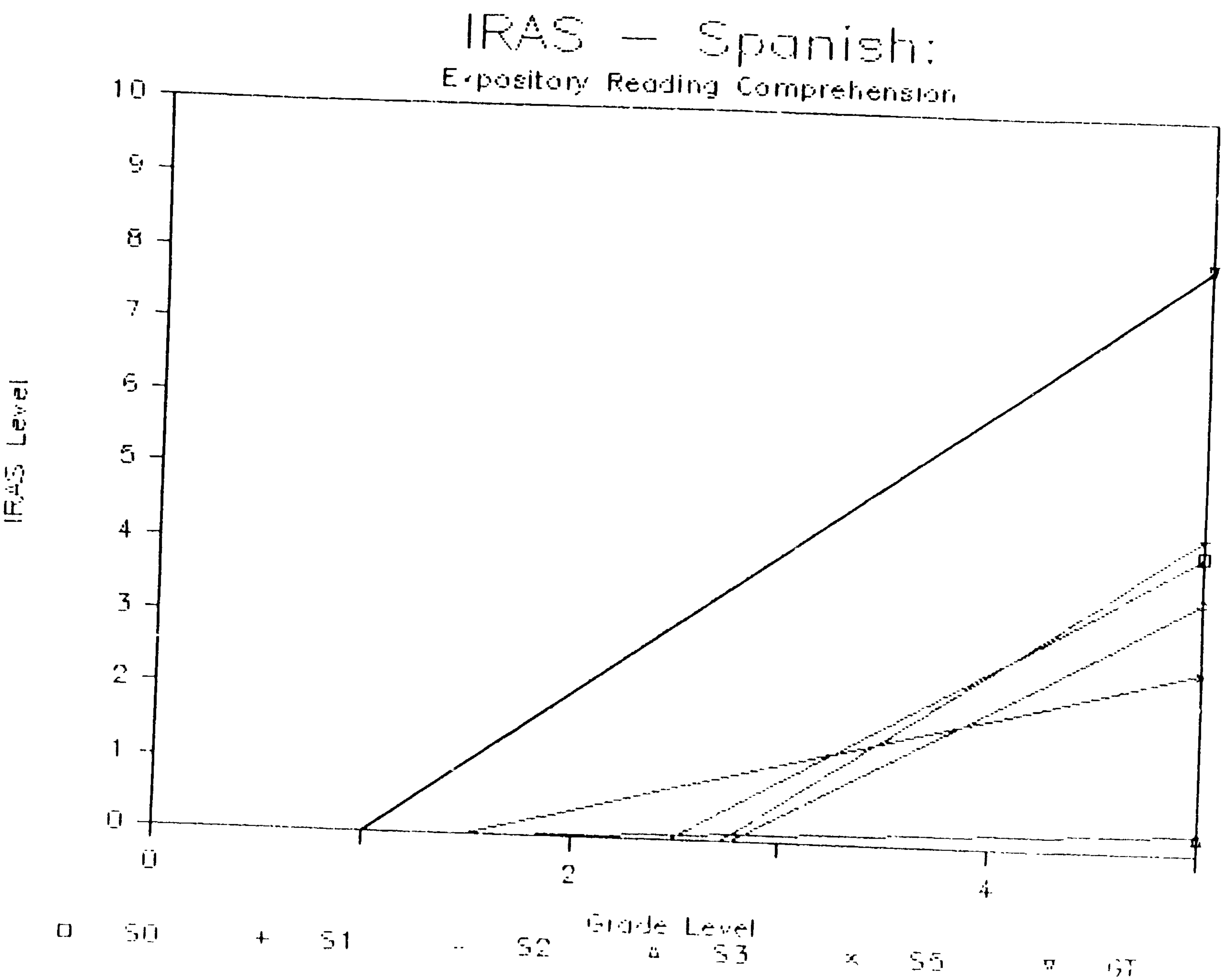


Figure 46. IRAS-S Expository Reading Comprehension growth for each site.

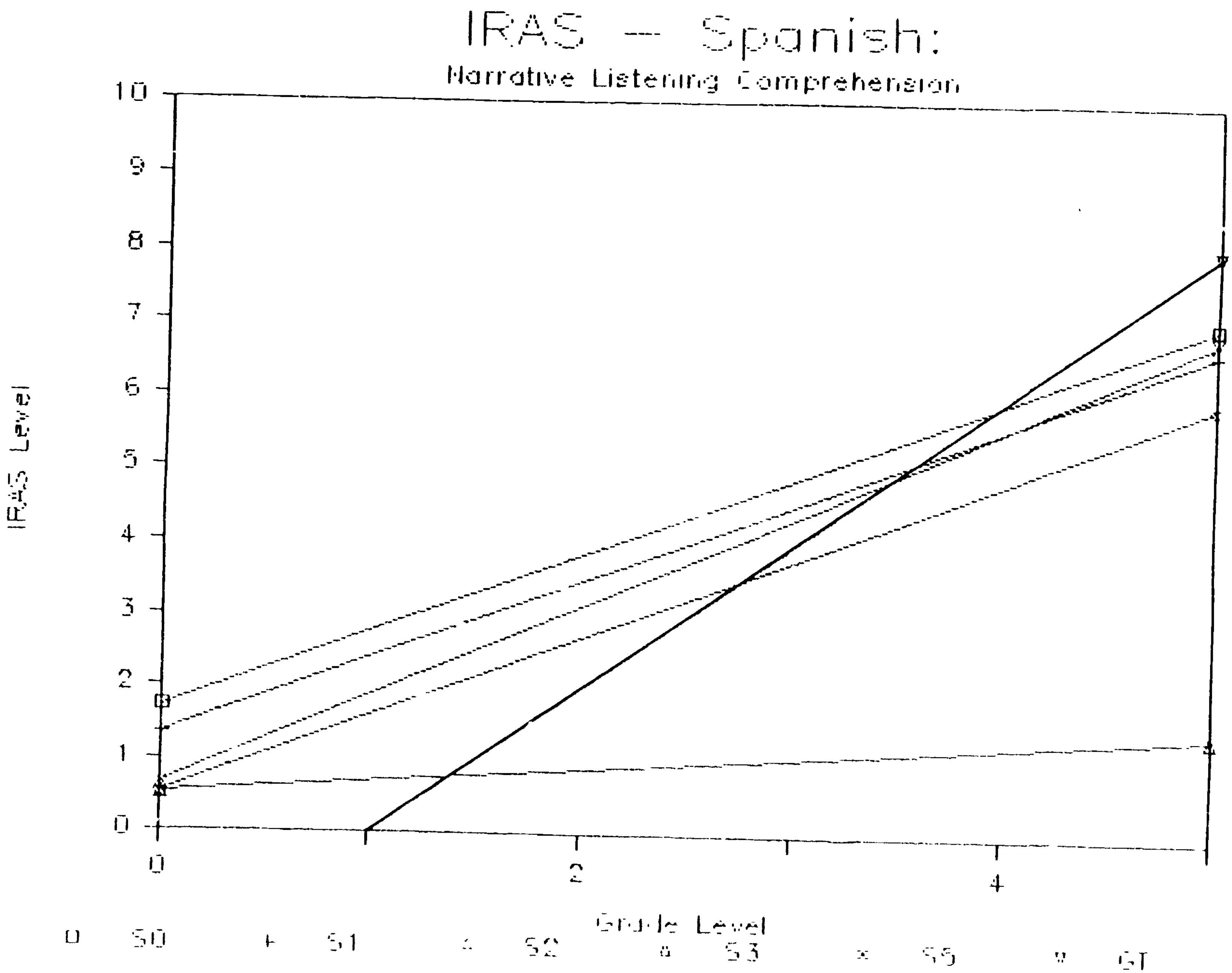


Figure 47. IRAS-S Narrative Listening Comprehension growth for each site.

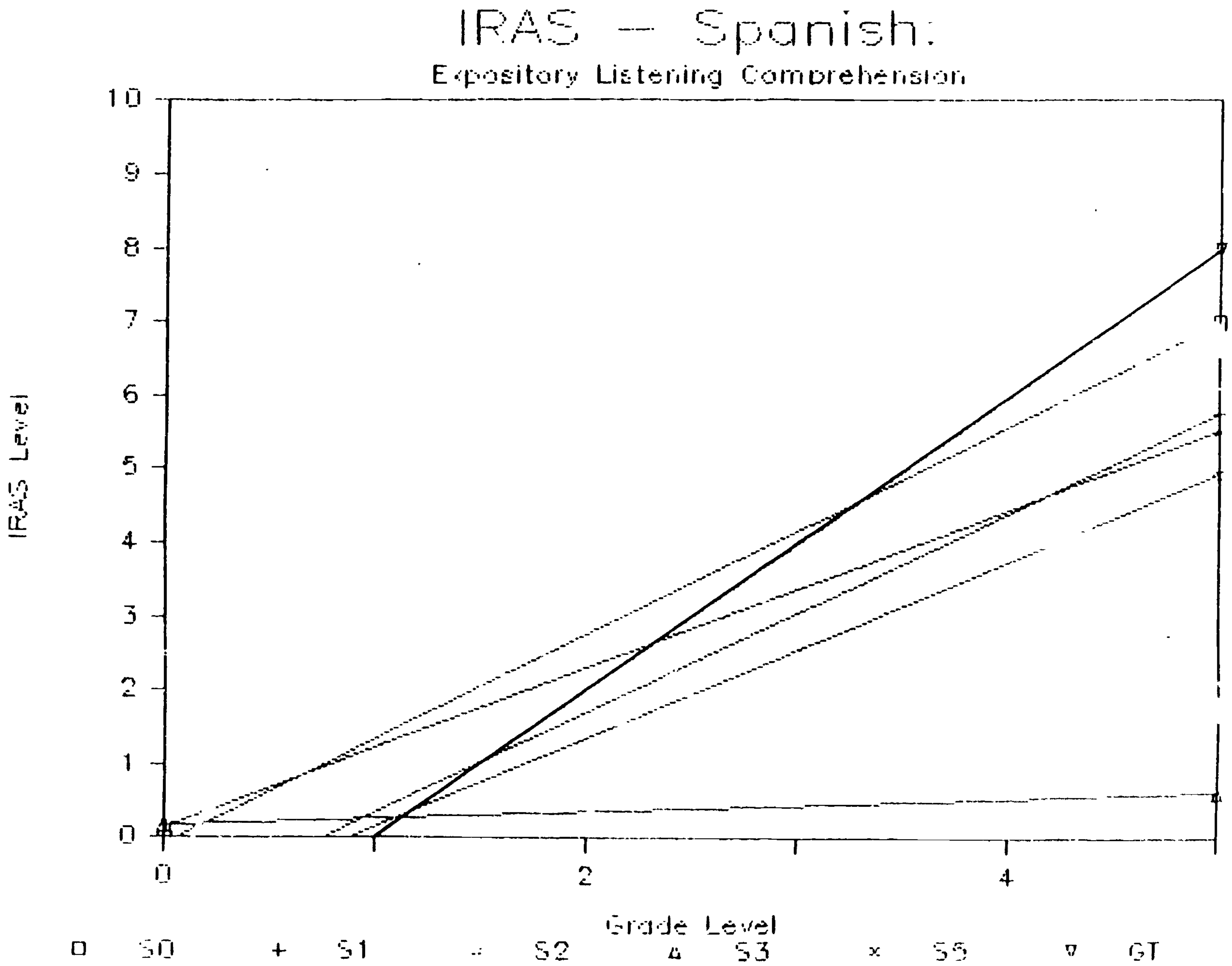


Figure 48. IRAS-S Expository Listening Comprehension growth for each site.

127

that for the non-border sites is substantially less for Site 3 (one level per year) and greater for Site 5 (2.6). Thus, for the site differences in word knowledge, all sites show growth above the growth track, except for Site 3, where performance is estimated to fall below expectation at second-grade entry. For Narrative and Expository Listening Comprehension (Figures 47 and 48), only the slope indices differ between sites -- the rate for Site 3 is only slightly above zero, while those for the remaining sites are close to one. Thus, for the formal language tasks, the aggregate picture generally holds for all sites except Site 3, where growth is much lower.

Site performance patterns in Real Word Decoding (Figure 40) are similar to those found in Vocabulary Definition. Sites 0 and 1 show relatively high student-intercepts (and low instruction-intercepts) with growth rates somewhat greater than expectation. Site 2 shows similar growth, but a low student-intercept (large instruction-intercept). Site 3 has a student-intercept at zero, with little subsequent growth, while Site 5 shows similar intercept values, but growth of 3 IRAS levels per year. The growth rates for the latter two sites are somewhat suspect given that they are based on only two data points; however, it is clear that Site 5 performance is substantially greater than that of Site 3. Synthetic Word Decoding (Figure 42) shows a similar pattern. Growth is comparable for the border sites with Site 2 showing a delay in the onset of effective instruction; growth at Site 5 is the highest (but also the least trustworthy estimate); and Site 3 shows little skill over the years of assessment. For Synthetic Word Spelling (Figure 43), only the slope indices differ by site: again, growth is comparable among the border sites, slightly greater at Site 5, and substantially lower (close to zero) at Site 3. Overall, the development of Spanish decoding skill is above expectations at Sites 0, 1, 2, and 5, though Site 2 shows a significant delay in the onset of effective instruction; at Site 3, there is little evidence of decoding skill acquisition.

For the reading tasks (Figures 44 through 46), the site differences are expressed as in the decoding tasks with two notable differences. First, in Narrative Reading Comprehension, both Sites 1 and 2 show a substantial delay in the onset of effective instruction, followed by growth which matches the model prediction. The remaining sites show earlier onsets of effective instruction (though still representing a delay from that predicted by the growth track), with growth rates at half the level of expectation for Sites 0 and 5, and close to zero for Site 3. A similar pattern holds for Expository Reading Comprehension, however, Site 0 now also shows the delay of instruction seen in Sites 1 and 2 for narrative texts.

Overall, Sites 0, 1, 2, and 5 fit the general aggregate picture given above: adequate growth in decoding skill (though instruction at Site 2 is delayed), but low growth in formal language skill, which results, by hypothesis, in low reading comprehension growth. Site 3, however, consistently shows low skill development in all of the IRAS tasks. A discussion in a subsequent volume of site differences in instructional programs will elucidate these findings.

Unexplained variance. The above discussion centered on the linear growth components revealed in the bilingual sample. However, an equally important issue concerns the nonlinearities found in the sample data, and these will be discussed here. The descriptive statistics for the bilingual sample for the average percent of unexplained variance for each of the IRAS scales are presented in Tables 35 and 36 (for English and Spanish, respectively). In these tables, the nine IRAS scales define the lefthand margin, and within each scale, the descriptive statistics of mean (M), standard deviation (S), and number of cases (N) are given. The tabled values are the percents of unexplained variance for each site under two conditions: no deletion and deletion. In the first case (no deletion), the percentages are based on all cases for which the best-fit line was computed. In the second case (deletion), the special cases of no-growth (100% unexplained variance) and two data points (0% unexplained variance) were removed. Thus, the change in the number of cases between the two treatments provides the total number of instances of these special cases. Since the growth functions for Sites 3 and 5 are all based on two data points, they have not been included under the Deletion heading. Further, for these two sites, the non-zero mean values tabled under No Deletion reflect the proportion of cases of no-growth (100% unexplained variance), as all other cases in these sites represent 0% unexplained variance.

For the English data (Table 35), the number of special cases is relatively small, except for Site 0, where, especially in the decoding skill tasks, the reduction is substantial (most of these are cases of no-growth). Most comparisons between the means under the deletion and no deletion treatments show little difference, indicating either few special cases (comparable Ns) or, where the differences in N are large, equal proportions of the two types of cases. Across sites, the percentages within each scale are relatively stable, with Vocabulary Definition, Synthetic Word Decoding, and Synthetic Word Spelling revealing the poorest fits at averages around 25%, the remaining scales around 15%. Note, however, that the standard deviations are substantial (generally of the same magnitude as the respective mean), indicating that although the aggregate measures show large linear components, many of the individual students do not.

For the Spanish data (Table 36), the fits are generally poorer. The average unexplained variance is around 25%, with the worst fits in Vocabulary Definition and Synthetic Word Spelling (38%). The reduction in the number of observations due to special cases is similar to that found in English for each site, with the exception of the Reading Comprehension scales for Site 2, which show a greater number of no-growth cases in Spanish. Again, note that the standard deviations are substantial.

Thus, while much of the variation in growth within both the English and Spanish scales can be said to be linear, there is, nonetheless, a significant amount of nonlinearity. As mentioned earlier, the predictions of these deviations from linearity by individual instructional variables will be the subject of a subsequent volume.

Table 35

Interactive Reading Assessment System - English:
Descriptive Statistics for the Bilingual Sample on the Percent of Unexplained Variance
for Each Scale for Each Site

		Percent of Unexplained Variance									
		No Deletion						Deletion			
Scale	Statistic	Overall	Site 0	Site 1	Site 2	Site 3	Site 5	Overall	Site 0	Site 1	Site 2
VDC	M	6.9	13.0	8.8	10.7	0.0	7.2	13.2	15.1	10.0	12.0
	S	17.6	18.2	12.7	10.8	0.0	26.1	15.4	18.8	13.1	10.7
	N	246	51	17	36	73	69	91	44	15	32
VDF	M	15.6	27.9	34.8	33.3	5.5	3.0	33.9	31.0	37.0	36.3
	S	28.6	28.8	28.9	35.5	22.9	17.1	31.0	28.7	28.4	35.5
	N	243	50	17	36	73	67	94	45	16	33
LDC	M	19.2	26.3	26.1	24.6	12.3	16.7	25.0	23.4	29.6	24.5
	S	33.4	31.4	17.5	32.7	33.1	37.6	23.0	20.6	15.5	28.9
	N	243	51	17	36	73	66	79	36	15	28
LSP	M	18.5	28.4	25.1	15.5	10.4	18.8	24.8	25.6	26.7	22.3
	S	33.3	34.4	25.7	22.1	30.8	39.4	25.3	26.9	25.7	23.5
	N	240	51	17	36	67	69	78	37	16	25
SRD	M	13.0	11.6	11.3	17.1	9.6	15.9	15.2	10.5	12.0	19.3
	S	30.4	27.6	10.4	28.6	29.6	36.9	22.2	9.8	10.3	29.7
	N	243	49	17	35	73	69	63	16	16	31
NRC	M	23.6	29.5	4.6	14.8	21.9	30.4	12.9	13.1	7.8	14.4
	S	39.3	39.0	6.5	21.6	41.7	46.4	14.0	13.0	6.8	16.5
	N	246	51	17	36	73	69	71	31	10	30
ERC	M	26.2	29.9	7.4	13.7	27.4	33.3	14.6	13.9	14.0	15.1
	S	41.8	43.9	11.2	20.0	44.9	47.5	14.6	19.2	12.1	13.5
	N	243	49	17	35	73	69	46	12	9	25
NLC	M	7.6	16.7	15.8	18.3	1.4	0.0	19.1	18.9	16.8	20.6
	S	15.5	20.1	14.1	16.1	11.7	0.0	17.8	20.4	13.9	15.7
	N	246	51	17	36	73	69	93	45	16	32
ELC	M	6.8	14.7	14.1	11.4	0.0	4.3	17.3	23.7	14.9	13.9
	S	19.0	27.9	13.7	16.6	0.0	20.5	20.7	27.3	13.7	17.4
	N	243	49	17	35	73	69	67	22	16	29

Table 36

Interactive Reading Assessment System - Spanish:
Descriptive Statistics for the Bilingual Sample on the Percent of Unexplained Variance
for Each Scale for Each Site

		Percent of Unexplained Variance									
		No Deletion						Deletion			
Scale	Statistic	Overall	Site 0	Site 1	Site 2	Site 3	Site 5	Overall	Site 0	Site 1	Site 2
VDC	M	25.3	25.0	23.0	16.8	48.0	5.8	24.9	26.7	27.9	20.2
	S	38.2	24.9	30.5	25.1	50.[illegible]	23.5	24.0	22.4	31.6	22.4
	N	248	51	17	36	75	69	83	44	14	25
VDF	M	21.7	42.8	33.9	29.9	12.3	8.7	37.7	44.3	36.0	26.[illegible]
	S	33.6	33.1	25.5	30.0	33.1	28.4	29.1	31.6	24.7	24.9
	N	245	51	17	35	73	69	93	47	16	30
LDC	M	40.9	32.7	27.6	30.6	78.7	12.5	29.8	31.4	29.3	28.1
	S	44.4	33.7	26.2	35.9	41.2	33.3	25.3	23.2	26.0	29.2
	N	243	51	17	36	75	64	75	34	16	25
LSP	M	46.0	36.6	28.4	40.9	90.1	14.[illegible]	37.9	36.5	30.2	43.8
	S	45.5	37.2	32.6	37.7	30.0	35.5	31.9	28.4	32.9	34.9
	N	244	51	17	36	71	69	77	32	16	27
SRD	M	33.5	17.0	11.5	19.5	78.7	8.7	14.9	16.8	12.3	15.3
	S	45.0	33.5	12.1	31.4	41.2	28.4	18.0	17.7	12.1	21.4
	N	245	49	17	35	75	69	55	14	16	25
NRC	M	65.0	49.5	14.9	37.5	97.3	68.1	17.3	17.7	7.5	22.6
	S	46.0	45.[illegible]	32.7	45.3	16.2	46.9	21.1	21.0	8.9	25.3
	N	248	51	17	36	75	69	42	24	7	11
EFL	M	66.9	45.9	30.4	43.7	97.3	69.6	15.7	16.3	2.8	22.9
	S	46.1	48.4	46.4	46.4	16.2	46.4	19.6	19.3	4.0	22.7
	N	245	49	17	35	75	69	25	9	6	10
NLC	M	20.5	23.9	29.8	29.3	26.7	4.3	29.9	25.9	31.7	31.9
	S	33.7	26.3	30.8	29.5	44.5	20.5	28.1	26.4	30.8	29.4
	N	248	51	17	36	75	69	96	47	16	33
ELC	M	30.9	13.9	29.7	33.8	57.3	13.0	30.5	12.2	25.3	34.9
	S	42.3	25.0	34.0	33.5	49.8	33.9	30.5	29.7	29.7	31.8
	N	245	49	17	35	75	69	71	24	16	31

Correlations

Having described the performance of the bilingual sample separately for the English and Spanish IRAS administrations, we now turn to a discussion of the correlations between scale performance. First, the correlations within language administrations will be discussed, then those found between language versions.

Within-language correlations. Table 37 displays the between scale correlations of the critical index measures for each instructional year. In each of the four panels, the English administration correlations are presented above the diagonal, with Spanish below. The correlation coefficients are generally larger in English than in Spanish, and both show steady increases in magnitude over the instructional years. The patterns of relations, however, remain fairly stable over the instructional years, and are quite similar for both English and Spanish.

The highest correlations for the formal language scales (Vocabulary Definition, and Narrative and Expository Listening Comprehension) are those obtained between each other. To a lesser degree, these scales are also related to the reading scales, where the relative magnitudes of the relations increase with each instructional year. Their weakest relations, though still not insubstantial, are with the decoding scales, which also grow in relative strength over the instructional years.

Similarly, for the decoding scales (Real Word Decoding, Synthetic Word Decoding, and Synthetic Word Spelling), the highest correlations at each instructional year are the intra-scale correlations between them. To a lesser degree, these scales are related to the reading scales, and these also show an increase in relative magnitude over the instructional years. Their weakest relations are with the formal language scales, again showing increases in relative magnitude with increases in schooling.

The reading scales (Sentence Reading, and Narrative and Expository Reading Comprehension) also follow this pattern: highest relations between each other, weaker relations with decoding, and weakest relations with formal language, with the latter two sets of relations increasing in relative magnitude with increases in instructional years.

Such a pattern in not unexpected -- the independence of the reading components assessed should be greatest in the early grades prior to the acquisition of skilled reading. However, with effective schooling, the relative differences in skill between component processes are reduced, and thus, the higher interrelatedness of the individual components at the later grades

The correlations between the growth measures computed over the critical indices for each scale are presented in Table 38. Again, the English administration values are presented above the diagonal, Spanish below. While the magnitudes of most of the coefficients are high

Table 37

Interactive Reading Assessment System - English and Spanish: Correlations Between Scale Critical Indices for Each Instructional Year for the Bilingual Sample

Instructional Year 1 (N=249)

	VDC CI	VDF CI	LDC CI	LSP PC	SRD SS	NRC CI	ERC CI	NLC CI	ELC CI
VDC CI	-	0.52	0.73	0.74	0.86	0.83	0.75	0.34	0.47
VDF CI	0.61	-	0.39	0.38	0.43	0.43	0.34	0.50	.50
LDC CI	0.79	0.64	-	0.77	0.69	0.60	0.55	0.38	0.45
LSP PC	0.76	0.60	0.92	-	0.64	0.61	0.59	0.31	0.38
SRD SS	0.79	0.62	0.76	0.74	-	0.88	0.79	0.36	0.47
NRC CI	0.61	0.41	0.53	0.54	0.57	-	0.89	0.37	0.48
ERC CI	0.42	0.25	0.35	0.38	0.27	0.69	-	0.36	0.49
NLC CI	0.31	0.58	0.39	0.37	0.39	0.25	0.18	-	0.80
ELC CI	0.36	0.57	0.47	0.44	0.39	0.33	0.23	0.88	-

Instructional Year 2 (N=248)

	VDC CI	VDF CI	LDC CI	LSP PC	SRD SS	NRC CI	ERC CI	NLC CI	ELC CI
VDC CI	-	0.75	0.76	0.77	0.78	0.77	0.73	0.53	0.63
VDF CI	0.68	-	0.58	0.59	0.60	0.62	0.58	0.61	0.64
LDC CI	0.85	0.58	-	0.72	0.66	0.64	0.60	0.43	0.51
LSP PC	0.87	0.63	0.89	-	0.63	0.64	0.60	0.45	0.51
SRD SS	0.86	0.61	0.77	0.78	-	0.86	0.83	0.53	0.66
NRC CI	0.77	0.49	0.64	0.67	0.82	-	0.95	0.54	0.75
ERC CI	0.74	0.44	0.57	0.59	0.78	0.96	-	0.50	0.75
NLC CI	0.52	0.63	0.53	0.51	0.49	0.49	0.44	-	0.86
ELC CI	0.53	0.60	0.51	0.49	0.53	0.53	0.51	0.88	-

Instructional Year 3 (N=93)

	VDC CI	VDF CI	LDC CI	LSP PC	SRD SS	NRC CI	ERC CI	NLC CI	ELC CI
VDC CI	-	0.77	0.75	0.74	0.77	0.77	0.76	0.64	0.68
VDF CI	0.72	-	0.53	0.62	0.59	0.62	0.59	0.53	0.54
LDC CI	0.82	0.57	-	0.70	0.69	0.68	0.67	0.48	0.58
LSP PC	0.87	0.59	0.82	-	0.64	0.60	0.62	0.52	0.56
SRD SS	0.80	0.56	0.69	0.74	-	0.87	0.88	0.59	0.72
NRC CI	0.76	0.58	0.64	0.64	0.80	-	0.95	0.79	0.81
ERC CI	0.77	0.56	0.61	0.62	0.79	0.96	-	.75	0.81
NLC CI	0.38	0.41	0.30	0.33	0.42	0.49	0.48	-	0.92
ELC CI	0.47	0.48	0.37	0.39	0.50	0.57	0.59	0.54	-

Instructional Year 4 (N=58)

	VDC CI	VDF CI	LDC CI	LSP PC	SRD SS	NRC CI	ERC CI	NLC CI	ELC CI
VDC CI	-	0.91	0.82	0.79	0.73	0.88	0.87	0.71	0.77
VDF CI	0.57	-	0.63	0.53	0.61	0.78	0.78	0.71	.7[illegible]
LDC CI	0.82	0.60	-	0.70	0.61	0.61	0.78	0.71	0.79
LSP PC	0.89	0.79	0.79	-	0.53	0.67	0.66	0.51	0.52
SRD SS	0.77	0.55	0.77	0.71	-	0.74	0.71	0.61	.54
NRC CI	0.75	0.62	0.71	0.74	0.80	-	0.96	0.84	.8[illegible]
ERC CI	0.78	0.66	0.69	0.75	0.78	0.97	-	0.79	[illegible]
NLC CI	0.51	0.57	0.36	0.47	0.74	0.57	0.59	-	.52
ELC CI	0.56	0.56	0.40	0.48	0.42	0.61	0.66	0.86	-

Note: English values are above the diagonal, Spanish below.

450

BEST COPY AVAILABLE

Table 78

Interactive Reading Assessment System - English and Spanish
Correlations Between Growth Indices for Each Scale
for the Bilingual Sample

	VDC SLOPE	VDC S-INT	VDC I-INT
VDC SLOPE	-	-0.49	0.36
VDC S-INT	-0.45	-	-0.81
VDC I-INT	0.29	-0.63	-

	VDF SLOPE	VDF S-INT	VDF I-INT
VDF SLOPE	-	-0.31	0.52
VDF S-INT	-0.68	-	-0.65
VDF I-INT	0.39	-0.67	-

	LDC SLOPE	LDC S-INT	LDC I-INT
LDC SLOPE	-	-0.72	0.47
LDC S-INT	-0.67	-	-0.77
LDC I-INT	0.41	-0.76	-

	LSP SLOPE	LSP S-INT	LSP I-INT
LSP SLOPE	-	-0.52	0.30
LSP S-INT	-0.51	-	-0.77
LSP I-INT	0.35	-0.72	-

	SRD SLOPE	SRD S-INT	SRD I-INT
SRD SLOPE	-	-0.59	0.42
SRD S-INT	-0.58	-	-0.74
SRD I-INT	0.48	-0.55	-

	NRC SLOPE	NRC S-INT	NRC I-INT
NRC SLOPE	-	-0.66	0.59
NRC S-INT	-0.78	-	-0.90
NRC I-INT	0.62	-0.85	-

	ERC SLOPE	ERC S-INT	ERC I-INT
ERC SLOPE	-	-0.68	0.59
ERC S-INT	-0.83	-	-0.82
ERC I-INT	0.77	-0.86	-

	NLC SLOPE	NLC S-INT	NLC I-INT
NLC SLOPE	-	-0.80	0.57
NLC S-INT	-0.51	-	-0.82
NLC I-INT	0.31	-0.72	-

	ELC SLOPE	ELC S-INT	ELC I-INT
ELC SLOPE	-	-0.76	0.54
ELC S-INT	-0.64	-	-0.78
ELC I-INT	0.37	-0.70	-

Note: English values are above the diagonal, Spanish below.

(absolute values ranging from .21 to .86), the three indices nonetheless, provide information that has a certain degree of independence, as seen in the descriptive accounts above.

The pattern of relations can be easily summarized. First, the correlations between slope and student-intercept are all negative, while those between slope and instruction-intercept are all positive: greater growth is associated with lower entry skill and greater delays in the onset of effective instruction. The relative magnitudes of the coefficients are generally greater for the relations between slope and student-intercept, as opposed to instruction-intercept, and results from the handling of the instruction-intercept special cases. Second, the relations between the two intercept measures are negative, and also generally represent the largest coefficients: greater entry skill is associated with earlier encountered effective instruction.

For both the English and Spanish administrations, the correlations between scales for individual growth indices are presented in Appendix F. These relations conform to those expected, given the relations found for the critical indices at each instructional year. Turning to the English data first, for the slope indices, growth is moderately related within the decoding scales (ranging from .35 to .50), to a lesser degree within the formal language scales (from .14 to .62), and to a greater degree within the reading scales (from .58 to .81). Further, growth in decoding is related to growth in reading (average value about .3), but not to growth in formal language; growth in formal language is, however, moderately related to growth in reading (about .35). For the student-intercept measures, again, the decoding scale entry points are related (ranging from .56 to .58), as are those for formal language (from .38 to .69), and for reading (from .48 to .77). Further, decoding entry is related to reading entry (average value about .4 for Sentence Reading, less for comprehension), and, to a lesser degree, to formal language entry (about .3); formal language and reading entry are also moderately related (about .3). The correlational pattern within the instruction-intercepts is similar to that found within the student-intercepts, except the coefficients are somewhat lower in magnitude.

For the Spanish data (also in Appendix F), the correlational pattern is very similar to that found in English, although there are some important differences. First, for the slope indices, the largest correlations are generally within the three sets of scales: formal language (.54 between the listening comprehension scales, but .1 for the relation between growth in definition skill and listening comprehension), decoding (ranging from .53 to .71), and reading (from .50 to .85). The intercorrelations are moderate between decoding and reading growth (average value of .4, though the relations between the synthetic word tasks and the comprehension tasks are much smaller), weaker between formal language and reading growth (about .2), and weakest between decoding and formal language (about .15). For the student-intercept measures, again, the decoding scale entry points are related (ranging from .66 to .76), as are those for formal language (from .27 to .56), and for reading (from .48 to .77). As in English, decoding entry is related to reading entry (average value about .4 for Sentence Reading, much less for the comprehension tasks), and, to a lesser degree, to formal language entry (about .4 for the definition task, .1 for the comprehension tasks); formal language and reading entry do not seem to be systematically related. As in English, the correlational pattern within the instruction-intercepts is similar to that found within the student-intercepts, except the coefficients are generally lower in magnitude.

As seen in these tables, the relationships between sl and student-intercept within a given scale are always negative: s the student-intercepts increase, the growth space decreases, and there is a trend for the rate of growth within that space to similarly decrease. However, looking at the relationships of these two growth indices between scales (tabled in Appendix F), some interesting relationships appear. First, for the English data, while the coefficients are generally negative, those between the student-intercepts of the decoding scales and the slopes of the reading comprehension scales are positive. While these English IRAS coefficients are not significantly different from zero, the corresponding coefficients for the Spanish IRAS are significant. Although the coefficients are moderate, this suggests that greater entry decoding skill is positively related to greater growth in subsequent reading comprehension. Further, within the Spanish IRAS data, there is also a trend, though not significant, for relatively greater formal language entry skill to be positively related to greater growth in reading comprehension.

Thus, for the within-language correlations of the IRAS formal language, decoding, and reading scales, the highest relationships are generally between the component scales within these three skill areas. The correlations between these areas are strongest for decoding and reading, somewhat weaker between formal language and reading, and weakest between decoding and formal language.

Between-language correlations. Table 39 displays the correlations between the English and Spanish critical indices for each instructional year. For the first instructional year, the strongest correlations (ranging from .40 to .53) are for the decoding scales along the diagonal: for a particular decoding task, initial skill in one language is related to initial skill in another. For the off-diagonal correlations, the highest values are within tasks requiring decoding skill: largest within the synthetic word tasks, next between the synthetic and real word tasks, and then between the decoding tasks and the reading tasks. For the latter relationships between decoding and reading, there is a marked trend for larger correlations between English decoding and Spanish readi g than between Spanish decoding and English reading. The formal language tasks (Vocabulary Definition, and Narrative and Expository Listening Comprehension) are, as expected, unrelated: for these young bilingual students, oral skill in one language is not rel d to oral skill in a second language.

A similar pattern for the second instructional year is found: the strongest relations tend to be for the decoding scales along the

Table 39

BEST COPY AVAILABLE

Interactive Reading Assessment System:
Correlations Between English and Spanish
Critical Indices for Each Instructional Year for the Bilingual Sample

Instructional Year 1 (N=250)

English \ Spanish	VDC CI	VDF CI	LDC CI	LSP PC	SRD SS	NRC CI	ERC CI	NLC CI	ELC CI
VDC CI	0.40	0.31	0.35	0.39	0.37	0.30	0.01	0.10	0.10
VDF CI	0.07	0.16	0.11	0.11	0.04	0.07	-0.06	-0.05	-0.05
LDC CI	0.48	0.38	0.53	0.52	0.49	0.30	0.11	0.18	0.21
LSP PC	0.41	0.30	0.42	0.46	0.39	0.23	0.04	0.16	0.16
SRD SS	0.27	0.22	0.23	0.28	0.40	0.06	-0.05	0.09	0.09
NRC CI	0.28	0.21	0.27	0.27	0.30	0.23	-0.08	0.04	0.01
ERC CI	0.20	0.17	0.22	0.22	0.35	-0.01	-0.07	0.11	0.06
NLC CI	0.08	0.12	0.14	0.09	0.06	0.02	-0.08	0.15	0.21
ELC CI	0.01	0.13	0.09	0.09	0.08	-0.11	-0.17	0.13	0.16

Instructional Year 2 (N=249)

English \ Spanish	VDC CI	VDF CI	LDC CI	LSP PC	SRD SS	NRC CI	ERC CI	NLC CI	ELC CI
VDC CI	0.53	0.30	0.45	0.47	0.51	0.47	0.45	0.12	0.18
VDF CI	0.32	0.24	0.27	0.30	0.32	0.31	0.27	0.04	0.12
LDC CI	0.49	0.24	0.48	0.47	0.51	0.40	0.38	0.14	0.19
LSP PC	0.39	0.22	0.39	0.40	0.42	0.28	0.29	0.09	0.15
SRD SS	0.23	0.10	0.14	0.18	0.37	0.30	0.32	-0.09	0.02
NRC CI	0.25	0.09	0.22	0.24	0.36	0.33	0.33	-0.03	0.02
ERC CI	0.20	0.06	0.16	0.16	0.36	0.27	0.30	-0.07	0.01
NLC CI	0.09	0.06	0.10	0.08	0.17	0.18	0.19	0.01	0.09
ELC CI	0.13	0.07	0.09	0.09	0.23	0.21	0.25	-0.03	0.06

Instructional Year 3 (N=95)

English \ Spanish	VDC CI	VDF CI	LDC CI	LSP PC	SRD SS	NRC CI	ERC CI	NLC CI	ELC CI
VDC CI	0.63	0.40	0.66	0.58	0.66	0.62	0.62	0.22	0.27
VDF CI	0.46	0.38	0.46	0.45	0.54	0.48	0.52	0.29	0.30
LDC CI	0.59	0.33	0.74	0.61	0.69	0.54	0.53	0.18	0.27
LSP PC	0.45	0.24	0.62	0.53	0.47	0.37	0.36	0.08	0.15
SRD SS	0.34	0.10	0.45	0.32	0.55	0.51	0.51	0.14	0.19
NRC CI	0.38	0.15	0.46	0.31	0.57	0.56	0.56	.26	0.25
ERC CI	0.38	0.16	0.46	0.29	0.56	0.55	0.56	0.30	0.24
NLC CI	0.20	0.10	0.27	0.18	0.38	0.36	0.35	0.41	[illegible]
ELC CI	0.27	0.14	0.31	0.21	0.43	0.43	0.46	0.35	0.35

Instructional Year 4 (N=58)

English \ Spanish	VDC CI	VDF CI	LDC CI	LSP PC	SRD SS	NRC CI	ERC CI	NLC CI	ELC CI
VDC CI	0.66	0.45	0.68	0.62	0.74	0.69	0.68	0.32	.35
VDF CI	0.31	0.41	0.51	0.50	0.59	0.58	0.61	.29	[illegible]
LDC CI	0.64	0.42	0.77	0.54	0.65	0.60	0.58	.36	.34
LSP PC	0.58	0.38	0.62	0.51	0.54	0.52	0.52	0.25	[illegible]
SRD SS	0.35	0.12	0.52	0.36	0.61	0.50	.46	[illegible]	[illegible]
NRC CI	0.55	0.26	0.62	0.49	0.73	0.65	0.62	.33	.35
ERC CI	0.55	.33	0.60	0.49	0.68	0.64	0.51	.34	[illegible]
NLC CI	0.45	0.30	0.53	0.42	0.60	[illegible]	0.57	.32	[illegible]
ELC CI	0.49	0.32	0.54	0.44	0.59	[illegible]	0.49	[illegible]	0.29

diagonal, followed by the off-diagonal correlations between decoding tasks. The reading scales are also interrelated, more so than found in the first instructional year. For the relations between decoding and reading, the relationships are, as earlier, stronger between the English decoding scales and the Spanish reading scales (ranging from .28 to .51) than between the Spanish decoding scales and the English reading scales (ranging from .06 to .25). Finally, as in the first instructional year, the relations of formal language skill between the two languages are minimal. Note, however, that the relative magnitudes of the relations between definition skill (especially in English) and decoding skill, across the two languages are increasing.

For the third and fourth instructional years, a slightly different pattern is found; its interpretation must be made cautiously since, given the cohort structure, the nature of this sample is substantially different from that involved in the first two instructional years. First, the diagonals no longer represent the highest correlation values; further, the correlations within decoding tasks and within reading tasks are not distinguishably higher than those between these sets of tasks. Thus, for the sample represented at these two instructional years, the decoding and reading tasks are all highly interrelated. Definition skill is again found to have relatively moderate correlations with the decoding and reading tasks, again, much more so for English definition skill than for Spanish. Further, the listening comprehension tasks begin to show moderate relations with the reading tasks, again, more so for English listening comprehension than for Spanish.

Thus, the correlational pattern between the yearly English and Spanish IRAS critical indices suggests a certain degree of transfer between decoding and reading skills; indeed, in later years, all of the decoding based tasks are interrelated across language. There is, however, a general trend for stronger relationships between a given English task across the set of Spanish tasks when compared to those relationships for the same given Spanish task across the set of English tasks -- literacy development in English seems to be more readily transferable to Spanish than vice versa.

The correlations for individual growth indices between English and Spanish scales are presented in Appendix F. First, note that the magnitude of the coefficients are relatively low in comparison to those that have just been discussed. While striking, the result is not unexpected. The growth functions provide information about trends in growth, and given that the linear fits of the functions are not perfect (see the above section on Unexplained variance), the relationships between the trends will be less sensitive to the yearly deviations that can be captured in the relationships within instructional years.

Considering first the relationships between the student-intercepts, the strongest relationships are within the decoding scales and reading scales, though those for the latter are not as widespread, and in fact, contain some negative coefficients (between English Sentence Reading and the two Spanish Reading Comprehension scales). Somewhat weaker relationships are found between the decoding and reading scales, more so for the Sentence Reading task than for the two Reading Comprehension tasks. Finally, small positive correlations appear between the formal language tasks. A similar pattern is found within the instruction-intercept correlations, though the relationships are not as widespread as those for the student-intercepts. For the slope indices, few of the correlations reach significance; the only systematic trend is between English Real Word Decoding and the Spanish tasks of Real Word Decoding and Reading Comprehension, though the coefficients are small.

Thus, the between-language correlations support the notion that, at least for English and Spanish, literacy skill developed in one language can be transferred to skill in a second language, especially for decoding and reading skills (less so for formal language skills). Further, the data also suggest that this transfer is, in some sense, more easily effected for English to Spanish than for Spanish to English. Concerning the latter, one interpretation, though by no means the only one, of the stronger relationship between initial skill in English decoding and Spanish reading compared to that between initial skill in Spanish decoding and English reading is as follows. If the acquisition of Spanish decoding skills is based on letter-sound correspondence rules which are (putatively) more regular than those in English, then their transfer to English reading skill may be more difficult, assuming sufficient English oral skill, than the transfer of initially acquired English decoding skills to Spanish reading, again, assuming sufficient Spanish oral skill. This may account for part of the transfer relationship between decoding and reading, but does not speak to the more generally found trend seen in the data and discussed above -- indeed, instructional differences in English and Spanish literacy may play a role in this trend, and these will be discussed in a subsequent volume.

Standardized Reading Achievement Tests

Generally beginning in first grade, standardized achievement tests are administered to all Texas students in the Spring of each year. In this study's sample of sites, three different standardized tests were used over the course of the data collection phase: the California Achievement Test (Sites 0, 1, and 2), the Comprehensive Test of Basic Skills (Site 3), and the Iowa Test of Basic Skills (Site 5). All standardized test data available for each target student, in the form of the total reading grade equivalent (as well as the date, form, and level of the test) were obtained from the schools each year.

Standardized achievement tests in Spanish were not administered systematically, nor to any great extent, by any of the schools in the study. Also, no such tests were administered in the monolingual-Spanish site (Site 4). As such, the few Spanish test scores that were collected do not provide a sufficient database to assess Spanish literacy growth, and are not discussed in this report.

The tasks, materials, scoring procedures, and psychometrics of the above standardized achievement tests are widely reported, and will not be treated here. Descriptive statistics for the bilingual sample are presented below, followed by a discussion of the correlations found between the growth indices computed.

Descriptive Statistics

In this section, the performance of the bilingual sample is described for the grade-level equivalents collected from the district-administered English standardized achievement tests. As noted previously in Volume 3, grade-level equivalents have been sharply criticized in the literature. However, for the arguments given earlier (also see Hoover, 1984), these values were employed as the primary index of standardized test performance.

The total reading grade equivalents were collected for each student during each year of participation in the study. As noted above, three different standardized tests were used in the participating sites, and in the analyses reported below, these differences in test have been ignored. Their analysis followed the same procedures employed in deriving the growth measures over the IRAS critical indices: linear growth functions were computed for each individual student over all available data points. The same caveats given for the IRAS growth functions also apply here -- most importantly, that the reliability of the growth estimates increases with the number of data points on which they are based. Thus, the estimates for Sites 0, 1, and 2 are the most trustworthy, as they are generally based on three to four values, while those from Sites 3 and 5 are least trustworthy, being derived from only two data points.

Table 40 displays the descriptive statistics for the three growth measures computed (slope, student-intercept, and instruction-intercept), giving the mean (M), standard deviation (S), and number of observations (N) for the performance of the sample overall, and by language entry category. First, note that multiple standardized test scores were only available for 30% of the sample, and thus, comparisons of standardized test performance with the other, more widely available literacy indices for the sample must be made cautiously. Second, the average unexplained variance, an index of the error in the best-fit linear functions, is 19.1 (with no-growth and two data point cases deleted) with a standard deviation of 23.1. Again, this indicates that although there is a substantial linear component represented in standardized test growth, there is also a certain degree of nonlinearity for individual students.

From Table 40, the average student-intercept of .9 suggests that, in the aggregate, the sample begins first grade just slightly below grade-level expectations. However, the instruction-intercept of -.7 (again, recalling the treatment of special cases) suggests that some literacy training began prior to school entry; given the global nature of the total reading grade equivalent, the component skills showing early instruction effects cannot be determined. The aggregate slope value of .8 suggests that the sample is not keeping a "year for year" pace between instruction and growth, and will be a full grade-level below expectation by fourth-grade exit.

Table 4[illegible]

Standardized Reading Achievement - English:
Descriptive Statistics on Growth Indices Overall
and by Language Entry Category for the Bilingual Sample

Scale	Measure	Statistic	Overall	Low Eng	High Eng	Low Span	High Span	Lo Lo	Lo Hi	Hi Lo	Hi Hi
SRAE	Slope	M	0.3	0.7	0.9	0.8	0.8	0.7	0.7	0.9	0.8
SRAE	Slope	S	0.6	0.6	0.6	0.7	0.5	0.7	0.4	0.7	0.5
SRAE	Slope	N	209	98	121	96	113	47	41	49	72
SRAE	S-Intrcp	M	0.8	0.7	1.1	0.9	0.9	0.8	0.5	1.1	1.2
SRAE	S-Intrcp	S	0.8	0.9	0.7	0.9	0.8	1.0	0.9	0.8	0.7
SRAE	I-Intrcp	M	-0.7	-0.3	-0.9	-0.6	-0.7	-0.4	-0.7	-0.9	-0.9
SRAE	I-Intrcp	S	1.8	1.7	1.8	1.8	1.8	1.7	1.8	1.9	1.8

458

These functions were subjected to an analysis of variance identical in structure to the analyses conducted on the IRAS growth indices (see discussion above for details). For the three growth measures analyzed, significant effects were only detected within the slope and student-intercept indices, and in both cases, only for the variables of English entry (low and high) and site (Sites 0, 1, 2, 3, and 5). Consequently, only these variables are discussed here, treating English entry first.

As mentioned above, significant differences were found for the English entry variable for both the slope ($F = 5.3$, $p < .02$) and student-intercept ($F = 4.5$, $p < .002$). Figure 49 presents the growth functions for the high and low English entry groups, along with the overall average growth function described earlier. For the low English entry group, the student-intercept is just below grade-level expectation, and growth proceeds at about .7 levels per year. For the high English entry group, the student-intercept is just above grade-level expectation, and growth is slightly higher (at .9) than that for the low English entry group. Thus, by fourth-grade exit, the high English group is projected to be about a half grade level behind, while the low English entry group is more than a full grade level lower.

Significant differences were also found for the variable of site, again, for both the slope ($F = 14.3$, $p < .001$) and student-intercept ($F = 5.7$, $p < .001$). Table 41 presents the descriptive statistics on the growth indices for the individual sites, both overall and by language entry category; and Figure 50 displays these site functions based on the tabled data. The figure suggests that performance is very similar for Sites 0, 2, and 3: student-intercepts slightly below grade-level expectation (more so for Site 2), coupled with slightly below expectation growth rates (.9). For Sites 1 and 5, student-intercepts are at or above grade-level expectation, but growth is much below expectation.

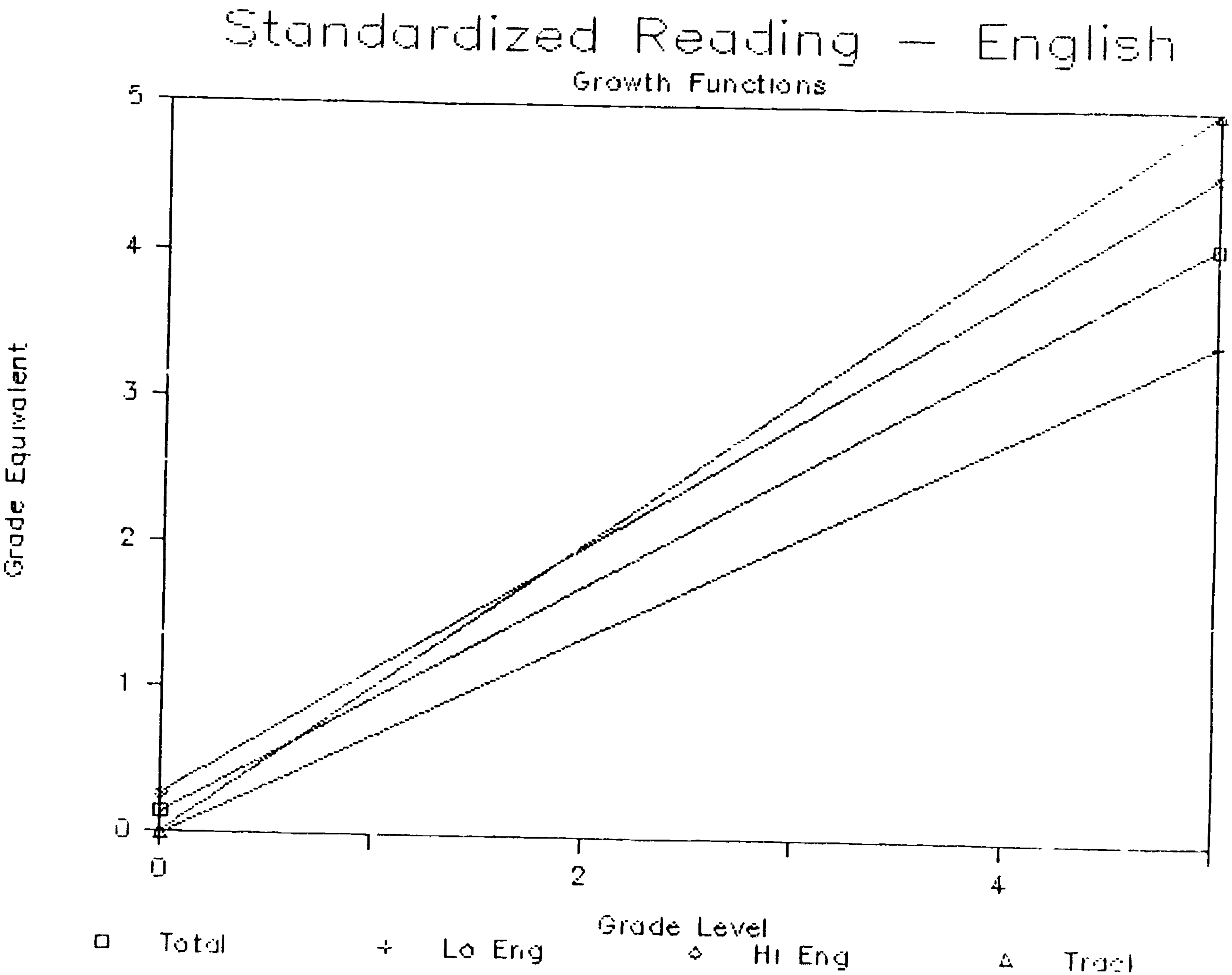
Thus, for the English standardized test indices, the aggregate picture is not substantially different from that found in the English IRAS reading comprehension tasks: student-intercept below expectation (somewhat more so for the IRAS indices), with subsequent growth close to expectation (slightly above in the IRAS indices, slightly below in the standardized test measures).

Correlations

The correlations between the three standardized test growth indices are as follows: slope and student-intercept, -.57, slope and instruction-intercept, .54, and between the two intercepts, -.62. Again, these values are very close to those found for the English IRAS reading comprehension growth indices, and will not be further discussed here.

Relationships Between Reading Achievement Measures

In earlier sections of this volume the measures derived from the central reading instruments employed in this study were discussed: both the English and Spanish versions of the Interactive Reading



143

Figure 49. Standardized Reading Achievement - English growth functions for the entire bilingual sample and by English entry category.

Table 4.

Standardized Reading Achievement - English
Descriptive Statistics on Growth Indices for Each Site
Overall and by Language Entry Category for the Bilingual Sample

Scale	Measure	Statistic	Site-0	Lo Lo-0	Lo Hi-0	Hi Lo-0	Hi Hi-0
SRAE	Slope	M	0.9	0.5	0.8		[illegible]
SRAE	Slope	S	0.3	0.3	[illegible]		[illegible]
SRAE	Slope	N	49	3	15		[illegible]
SRAE	S-Intrcp	M	[illegible]	0.7	0.6		[illegible]
SRAE	S-Intrcp	S	0.7	0.5	[illegible]		[illegible]
SRAE	I-Intrcp	M	-0.6	-1.3	-0.1		-0.7
SRAE	I-Intrcp	S	1.5	2.2	1.2		[illegible]
			Site-1	Lo Lo-1	Lo Hi-1	Hi Lo-1	Hi Hi-1
SRAE	Slope	M	0.7	0.8	0.7	0.6	0.6
SRAE	Slope	S	0.3	0.4	0.3	0.2	[illegible]
SRAE	Slope	N	17	4	6	[illegible]	3
SRAE	S-Intrcp	M	1.0	0.5	0.8	1.5	[illegible]
SRAE	S-Intrcp	S	0.6	0.2	0.6	0.7	0.5
SRAE	I-Intrcp	M	-1.0	0.2	-1.1	-1.6	-1.9
SRAE	I-Intrcp	S	1.9	0.6	2.3	1.7	2.7
			Site-2	Lo Lo-2	Lo Hi-2	Hi Lo-2	Hi Hi-2
SRAE	Slope	M	0.9	0.7	0.7	[illegible]	[illegible]
SRAE	Slope	S	0.4	0.1	0.4	0.4	[illegible]
SRAE	Slope	N	35	[illegible]	15	4	[illegible]
SRAE	S-Intrcp	M	[illegible]	[illegible]	0.1	[illegible]	[illegible]
SRAE	S-Intrcp	S	[illegible]	0.8	0.8	[illegible]	[illegible]
SRAE	I-Intrcp	M	[illegible]	0.7	[illegible]	[illegible]	[illegible]
SRAE	I-Intrcp	S	1.5	[illegible]	2.2	[illegible]	[illegible]
			Site-3	Lo Lo-3	Lo Hi-3	Hi Lo-3	Hi Hi-3
SRAE	Slope	M	0.9	0.8		[illegible]	[illegible]
SRAE	Slope	S	0.7	0.7		[illegible]	[illegible]
SRAE	Slope	N	[illegible]	26		[illegible]	[illegible]
SRAE	S-Intrcp	M	0.8	0.6		0.9	[illegible]
SRAE	S-Intrcp	S	0.7	0.7		[illegible]	[illegible]
SRAE	I-Intrcp	M	-0.5	[illegible]		-0.5	[illegible]
SRAE	I-Intrcp	S	1.8	1.7		[illegible]	[illegible]
			Site-5	Lo Lo-5	Lo Hi-5	Hi Lo-5	Hi Hi-5
SRAE	Slope	M	0.5	[illegible]	0.4	[illegible]	[illegible]
SRAE	Slope	S	0.7	0.9	0.6	[illegible]	[illegible]
SRAE	Slope	N	41	6	[illegible]	[illegible]	[illegible]
SRAE	S-Intrcp	M	1.4	[illegible]	[illegible]	[illegible]	[illegible]
SRAE	S-Intrcp	S	1.1	[illegible]	[illegible]	[illegible]	[illegible]
SRAE	I-Intrcp	M	-1.4	-0.1	[illegible]	[illegible]	[illegible]
SRAE	I-Intrcp	S	[illegible]	[illegible]	1.4	[illegible]	[illegible]

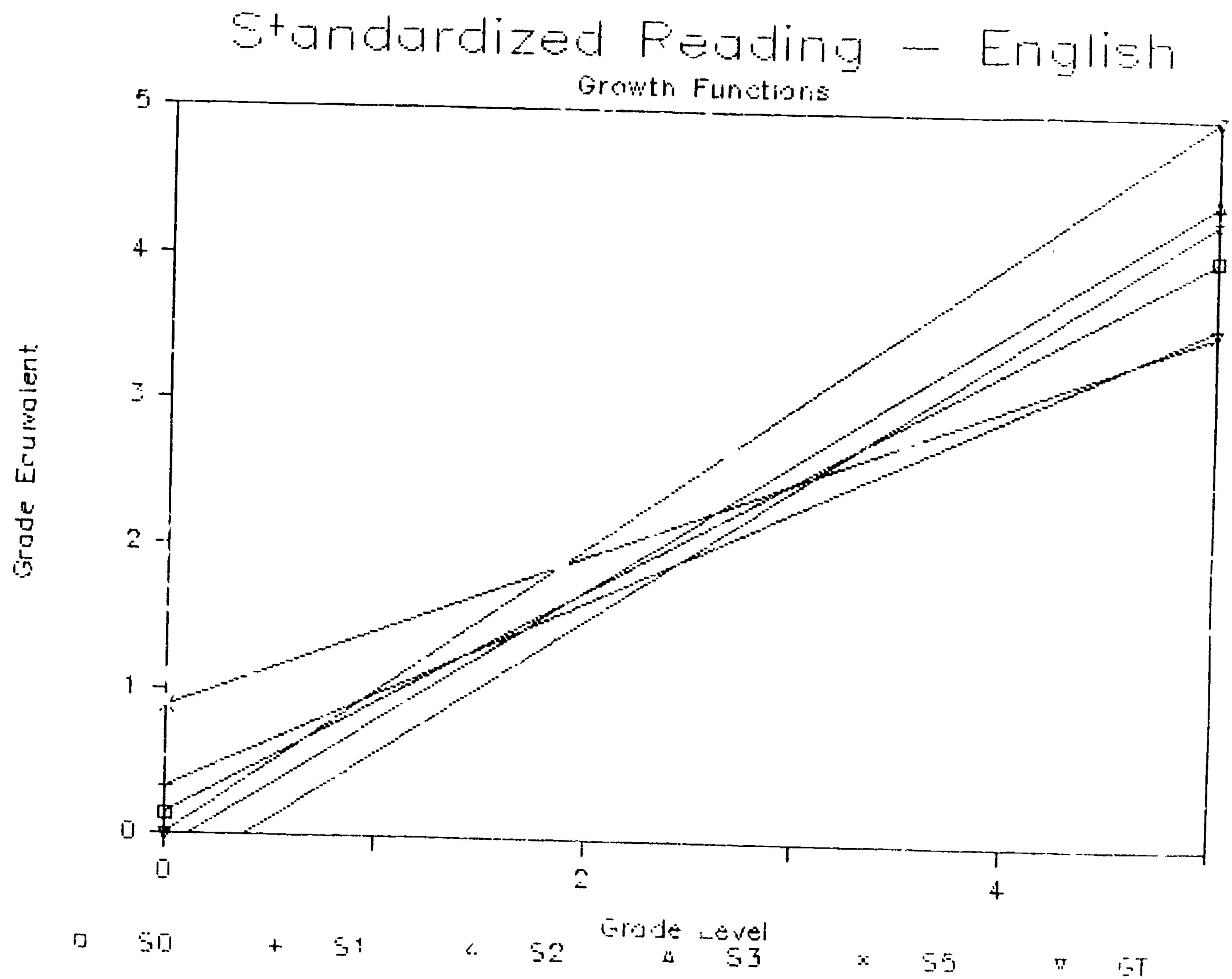


Figure 50. Standardized Reading Achievement - English growth functions for each site for the bilingual sample.

Assessment System, and the English standardized reading achievement tests. Here, the discussion will focus on the correlations obtained between the measures derived from these instruments. Since the relationships between the English and Spanish IRAS administrations were presented above in the discussion of that instrument, this section will focus on the relationships between the IRAS summary measures and the English standardized reading achievement test measures.

Table 42 presents the correlations between the English IRAS critical indices and the standardized reading achievement grade equivalents for each instructional year. For the first instructional year (top panel), the correlations are highest for the IRAS decoding and reading scales (ranging from .52 to .66), and lowest for the IRAS formal language scales (ranging from .27 to .32). In the remaining instructional years (bottom three panels), all coefficients are generally comparable, but of greater magnitude (ranging from .47 to .75). Keeping in mind the difference in samples between the first two instructional years and the last two instructional years, it seems that decoding and reading skills largely drive standar ed test performance in the initial stages of reading acquisition. However, as reading acquisition progresses, the degree of interrelatedness between component skills increases, as was seen in the IRAS scale correlations discussed above, and all of the IRAS scales show stronger relations with the standardized grade equivalents.

Table 43 displays the correlations between the English IRAS and standardized reading achievement growth indices. The coefficients are not as large as those just seen for the yearly indices, again, for the same reasons discussed above for the IRAS between scale correlations. The most systematic correlations between the two instruments are between slope indices, between student-intercepts, and between IRAS instruction-intercepts and standardized reading achievement student-intercepts. For the slope indices, significant, but generally small, correlations are found with the IRAS decoding and reading scales (ranging from .14 to .32). For the student-intercepts, modera e correlations are found for all of the IRAS scales (ranging from .24 to .47). The instruction-intercepts are not correlated between the two instruments, but all of the instruction-intercepts for the IRAS scales (except for Vocabulary Definition) are moderately (and negatively) correlated with the standardized reading achievement student-intercepts (ranging from -.42 to -.22).

The correlations between the Spanish IRAS and the English standardized reading achievement growth indices are presented in Table 44. The only significant correlations found are between the student-intercepts and concern only the IRAS decoding scales (with positive coefficients). This provides further support to the notion that decoding skills can be transferred between English and Spanish.

Thus, for the moderate correlations between the English reading growth indices (a) growth in decoding is positively related to growth in standardized reading achievement performance, (b) student skill at entry on all IRAS component skill assessments is positively related to standardized reading test entry, and (c) the onset of effective

BEST COPY AVAILABLE

Table 41

Reading Measures - English:
Correlations Between IRAS Critical Indices and Standardized Reading Achievement
Grade Equivalents for Each Instructional Year for the Bilingual Sample

Instructional Year 1 (N=190)
English

	VDC CI	VDF CI	LDC CI	LSP PC	SRD SS	NRC CI	ERC CI	NLC CI	ELC CI
SRD EQV	0.66	0.32	0.52	0.65	0.59	0.63	0.55	0.27	0.32

Instructional Year 2 (N=229)
English

	VDC CI	VDF CI	LDC CI	LSP PC	SRD SS	NRC CI	ERC CI	NLC CI	ELC CI
SRD EQV	0.59	0.60	0.54	0.55	0.73	0.74	0.75	0.52	.63

Instructional Year 3 (N=93)
English

	VDC CI	VDF CI	LDC CI	LSP PC	SRD SS	NRC CI	ERC CI	NLC CI	ELC CI
SRD EQV	0.63	0.47	0.54	0.52	0.62	0.69	0.69	0.62	0.61

Instructional Year 4 (N=58)
English

	VDC CI	VDF CI	LDC CI	LSP PC	SRD SS	NRC CI	ERC CI	NLC CI	ELC CI
SRD EQV	0.69	0.60	0.59	0.49	0.62	0.68	0.66	0.56	0.56

Table 4[illegible]

Reading Measures - English:
Correlations Between IRAS and Standardized Reading Achievement
Growth Indices for the Bilingual Sample

IRAS - English (N=205)

Slope

	VDC SLOPE	VDF SLOPE	LDC SLOPE	LSP SLOPE	SRD SLOPE	NRC SLOPE	ERC SLOPE	NLC SLOPE	ELC SLOPE
SEQ SLOPE	0.25	0.04	0.14	0.32	0.22	0.18	0.15	-0.07	-0.03
SEQ S-INT	[illegible]	-0.02	-0.09	-0.01	-0.08	-0.11	-0.04	-0.03	[illegible]
SEQ I-INT	0.10	0.03	0.23	0.07	0.07	0.09	0.00	0.05	0.14

S-Intercept

	VDC S-INT	VDF S-INT	LDC S-INT	LSP S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
SEQ SLOPE	0.03	0.13	.00	-0.19	0.10	0.08	0.08	0.12	0.12
SEQ S-INT	0.45	0.24	0.35	0.47	0.33	0.36	0.29	0.25	0.28
SEQ I-INT	-0.17	-0.08	-0.07	-0.19	-0.07	-0.13	-0.05	-0.1[illegible]	-0.15

I-Intercept

	VDC I-INT	VDF I-INT	LDC I-INT	LSP I-INT	SRD I-INT	NRC I-INT	ERC I-INT	NLC I-INT	ELC I-INT
SEQ SLOPE	0.01	-0.12	-0.01	0.02	-0.09	-0.02	-0.03	-0.10	-0.16
SEQ S-INT	-0.42	-0.12	-0.29	-0.28	-0.24	-0.39	-0.31	-0.22	-0.3[illegible]
SEQ I-INT	.13	-0.01	0.19	0.11	.00	0.17	0.16	0.15	0.11

BEST COPY AVAILABLE

467

Table 44

Reading Measures:
Correlations Between IRAS - Spanish and Standardized Reading
Achievement - English Growth Indices for the Bilingual Sample

IRAS - Spanish (N=200)

Slope

	VDC SLOPE	VDF SLOPE	LDC SLOPE	LSP SLOPE	SRD SLOPE	NRC SLOPE	ERC SLOPE	NLC SLOPE	ELC SLOPE
GEQ SLOPE	[illegible]	-0.14	-0.06	0.02	0.07	0.05	0.04	0.05	-0.01
GEQ S-INT	0.09	0.16	0.00	-0.07	0.05	[illegible]	0.18	[illegible]	0.12
GEQ I-INT	0.11	-0.07	0.07	0.03	0.02	-0.01	-0.03	0.01	-0.09

S-Intercept

	VDC S-INT	VDF S-INT	LDC S-INT	LSP S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
GEQ SLOPE	-0.11	0.05	0.06	-0.06	-0.04	-0.06	-0.05	-0.05	[illegible]
GEQ S-INT	0.22	0.06	0.22	0.33	0.16	0.01	-0.06	0.06	-0.02
GEQ I-INT	-0.18	-0.04	-0.13	-0.07	-0.06	-0.02	.00	-0.07	0.01

I-Intercept

	VDC I-INT	VDF I-INT	LDC I-INT	LSP I-INT	SRD I-INT	NRC I-INT	ERC I-INT	NLC I-INT	ELC I-INT
GEQ SLOPE	0.12	-0.03	-0.03	-0.05	-0.01	0.04	0.05	0.04	0.01
GEQ S-INT	-0.24	-0.10	-0.19	-0.16	-0.08	-0.01	0.04	.00	[illegible]
GEQ I-INT	0.13	0.04	0.07	-0.02	.00	0.06	0.04	.00	[illegible]

instruction for each of the IRAS components is negatively related to student standardized test skill at first-grade entry. For the Spanish-English correlations discussed, the only notable relationships are between Spanish IRAS decoding entry skill and English standardized reading achievement entry skill. We now turn to a discussion of the relations between the pre-reading and reading measures.

Relationships Between Pre-reading and Reading Achievement Measures

In this section, the correlations between the pre-reading summary measures (derived from the English and Spanish administrations of the Stanford Foundation Skills Test) and the reading measures (derived from the English and Spanish administrations of the Interactive Reading Assessment System, and the English standardized reading achievement administrations) are discussed.

Tables 45 and 46 present the correlations between the English SFST summary measures and the English IRAS growth indices for the kindergarten and first-grade bilingual samples, respectively. For both tables, the SFST correlations with the IRAS slope indices are displayed in the top panels, with the student-intercept indices in the middle panels, and with the instruction-intercept indices in the bottom panels. Given the small, and unrepresentative, first-grade sample (see the discussion of the SFST above for its composition), only the data from the more representative kindergarten sample will be discussed here.

Considering the student-intercept correlations for this sample first (Table 45, middle panel), alphabet knowledge shows moderate correlations (ranging from .20 to .49) across the IRAS scales, again revealing its strength as a general predictor of early literacy skill. Word naming also shows moderate correlations with the decoding and reading indices (ranging from .29 to .39), but not generally (as would be expected) with the formal language scales. The two letter matching tasks show somewhat smaller correlations, and only with the IRAS tasks of Real Word Decoding and Sentence Reading (ranging from .22 to .31) -- such visual skills would be expected to be related to tasks requiring decoding, over those tasks which are completely oral. The transfer task of Auditory-Phonetic Segmentation shows an expected correlation with Synthetic Word Decoding (.22), but, is also found to be moderately related to the formal language tasks. The latter relation suggests that the metalinguistic skill assessed by the segmentation task is more advanced (or at least more accessible) in students with stronger oral language skill. Finally, the SFST oral language tasks of Vocabulary, Definition, and Comprehension show moderate relationships to each of the formal language tasks, and to Real Word Decoding.

For the instruction-intercepts (Table 45, bottom panel), again, alphabet knowledge reveals a general predictive capacity, showing moderate, negative correlations with all of the IRAS indices (except for Expository Reading Comprehension), ranging in value from -.22 to -.41 -- early acquisition of the alphabet is associated with earlier onsets of effective instruction in each of the IRAS components. The remaining scales reveal patterns consistent with those found in the

Table 45

Pre-reading and Reading Measures - English:
Correlations Between IRAS Growth Indices and
SFST Summary Measures for the Kindergarten Bilingual Sample

IRAS - English (N=157)

Slope

SFST English Kinder	VDC SLOPE	VDF SLOPE	LDC SLOPE	LSP SLOPE	SRD SLOPE	NRC SLOPE	ERC SLOPE	NLC SLOPE	ELC SLOPE
ALPHPR	0.16	-0.14	-0.02	0.25	0.16	0.18	0.20	-0.11	0.01
WRDNAM	0.12	-0.22	-0.04	0.11	0.04	0.07	0.15	-0.04	0.04
SNGLTM	0.01	-0.05	-0.12	0.08	-0.10	-0.01	.00	-0.10	-0.08
DBLLTM	0.01	0.01	-0.14	0.10	-0.12	-0.01	0.01	-0.06	-0.05
PSTRNG	0.10	-0.01	-0.01	0.13	0.12	0.07	0.09	-0.01	0.06
PSTRNF	0.20	-0.04	0.08	0.20	0.13	0.14	0.17	-0.18	-0.08
VCFDCL	0.03	-0.21	-0.13	0.17	0.22	0.19	0.22	-0.15	-0.10
DEFNTN	-0.06	-0.32	-0.12	0.03	-0.01	0.11	0.24	-0.17	-0.32
CMPPRE	0.04	-0.21	-0.12	0.16	0.10	0.13	0.21	-0.29	-0.24

S-Intercept

SFST English Kinder	VDC S-INT	VDF S-INT	LDC S-INT	LSP S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
ALPHPR	0.49	0.35	0.41	0.36	0.39	0.32	0.20	0.32	0.29
WRDNAM	0.29	0.33	0.31	0.18	0.39	0.31	0.15	0.18	.18
SNGLTM	0.22	0.16	0.19	0.14	0.24	0.16	0.17	0.14	0.15
DBLLTM	0.23	0.11	0.20	0.12	0.31	0.16	0.14	0.20	0.18
PSTRNG	0.12	0.19	0.19	0.13	0.06	0.13	0.10	0.21	0.15
PSTRNF	0.15	0.25	0.21	0.10	0.15	0.12	0.09	0.35	0.31
VCFDCL	[illegible]	0.42	0.31	0.18	0.13	0.11	0.11	.46	0.42
DEFNTN	0.34	.43	0.33	0.29	0.24	0.12	.30	.35	0.27
CMPPRE	[illegible]	.33	0.25	0.13	0.12	0.07	0.01	0.50	0.48

I-Intercept

SFST English Kinder	VDC I-INT	VDF I-INT	LDC I-INT	LSP I-INT	SRD I-INT	NRC I-INT	ERC I-INT	NLC I-INT	ELC I-INT
ALPHPR	-0.41	-0.25	-0.37	-0.22	-0.33	-0.25	-0.14	-0.33	-0.27
WRDNAM	-0.13	0.02	-0.31	-0.10	-0.24	-0.20	-0.07	-0.24	-0.15
SNGLTM	-0.09	-0.04	-0.24	-0.10	-0.20	-0.10	-0.11	-0.15	-0.22
DBLLTM	-0.17	-0.06	-0.18	.07	-0.24	-0.15	-0.11	-0.25	[illegible]
PSTRNG	-0.12	-0.08	-0.07	-0.07	-0.13	-0.12	[illegible]	-0.22	-0.14
PSTRNF	-0.15	-0.18	0.05	-0.06	-0.07	-0.05	-0.08	-0.30	-0.22
VCFDCL	-0.29	-0.24	-0.12	-0.12	-0.11	-0.08	-0.06	-0.46	-0.35
DEFNTN	-0.23	-0.19	-0.21	-0.16	-0.15	-0.14	-0.02	-0.39	[illegible]
CMPPRE	[illegible]	-0.23	-0.21	-0.05	-0.09	-0.11	[illegible]	[illegible]	-0.38

Table 45

Pre-reading and Reading Measures - English:
Correlations Between IRAS Growth Indices and
SFST Summary Measures for the First-grade Bilingual Sample

IRAS - English (N=49)

Slope

		VDC SLOPE	VDF SLOPE	LDC SLOPE	LSP SLOPE	SRD SLOPE	NRC SLOPE	ERC SLOPE	NLC SLOPE	ELC SLOPE
	ALPHPR	0.25	-0.25	-0.04	0.15	0.06	-0.16	-0.41	-0.42	-0.48
	WRDNAM	-	-	-	-	-	-	-	-	-
	ENGLTM	0.03	-0.15	0.04	-0.02	-0.05	-0.18	-0.10	-0.08	-0.13
SFST	DBLLTM	0.25	-0.25	0.16	-0.01	-0.14	-0.06	-0.14	-0.19	-0.23
English	PSTRNG	0.17	-0.02	-0.01	0.03	0.07	0.22	0.10	0.17	0.22
First	PSTRNF	0.21	-0.30	0.05	0.14	-0.28	-0.09	0.03	-0.03	-0.11
	VCFDCL	0.27	-0.28	0.11	0.20	0.05	0.06	-0.05	-0.05	-0.30
	DEFNTN	-	-	-	-	-	-	-	-	-
	IMPPRB	-	-	-	-	-	-	-	-	-

S-Intercept

		VDC S-INT	VDF S-INT	LDC S-INT	LSP S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
	ALPHPR	0.60	0.58	0.51	0.47	0.45	0.45	0.57	0.61	0.63
	WRDNAM	-	-	-	-	-	-	-	-	-
	ENGLTM	0.27	0.29	0.09	0.24	0.22	0.28	0.20	0.38	0.25
SFST	DBLLTM	0.29	0.51	0.17	0.46	0.27	0.19	0.25	.48	0.34
English	PSTRNG	0.05	0.11	0.21	0.09	0.14	-0.05	-0.01	0.05	-0.04
First	PSTRNF	0.27	0.48	0.27	0.33	0.41	0.24	[illegible]	0.44	0.26
	VCFDCL	0.59	0.71	0.35	0.32	0.49	0.37	0.33	0.65	0.60
	DEFNTN	-	-	-	-	-	-	-	-	
	IMPPRB	-	-	-	-	-	-	-	-	-

I-Intercept

		VDC S-INT	VDF S-INT	LDC S-INT	LSP S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
	ALPHPR	-0.62	-0.44	-0.36	-0.41	-0.21	-0.41	-0.55	-0.51	-0.62
	WRDNAM	-	-	-	-	-	-	-	-	-
	ENGLTM	-0.24	-0.09	-0.03	-0.18	[illegible]	-0.23	-0.29	-0.33	-0.28
SFST	DBLLTM	-0.29	-0.31	-0.16	-0.33	[illegible]	-0.34	-0.34	-0.37	[illegible]
English	PSTRNG	-0.11	-0.15	-0.19	-0.03	-0.12	0.23	0.07	-0.10	[illegible]
First	PSTRNF	-0.27	-0.31	-0.23	-0.25	-0.27	-0.20	-0.19	-0.41	-0.24
	VCFDCL	-0.65	-0.63	-0.21	-0.28	-0.32	-0.37	-0.42	-0.55	-0.62
	DEFNTN	-	-	-	-		-	-	-	-
	IMPPRB	-	-	-	-	-	-	-	-	-

BEST COPY AVAILABLE

student-intercepts, however, these relations are neither as widespread, nor as systematic.

For the slope correlations (Table 45, top panel), some interesting patterns appear. One may consider the SFST measures as entry values (similar to student-intercepts at kindergarten), and as such, may expect them to behave as the previously discussed correlations between slope and student-intercept, which, for all of the significant correlations between these indices in the English IRAS, are negative. Here, however, many of the significant correlations are positive: alphabet knowledge with Synthetic Word Spelling and Expository Reading Comprehension; segmentation skill with Synthetic Word Spelling; the set of oral language SFST indices with the reading comprehension measures. These relations indicate, for the appropriate scales, that not only is higher skill at kindergarten entry positively related to higher skill at first-grade entry, as seen above, but it is also related to relatively higher average growth after first-grade entry. Nonetheless, some correlations show the usual negative relationships, as between the SFST oral language tasks and the IRAS formal language tasks.

The correlations between the Spanish SFST summary measures and the Spanish IRAS growth indices are displayed in Tables 47 and 48 for the kindergarten and first-grade bilingual samples, respectively. These tables are organized in the same fashion as those just discussed; again, only the kindergarten data will be treated here. First, for the student-intercepts (Table 47, middle panel), alphabet knowledge is related only to Synthetic Word Spelling (.29); recalling that Spanish alphabet knowledge is minimal in the sample, its poor predictive power is expected. Second, the oral language tasks for the two instruments are related, ranging from .26 to .57, and indicate that relative oral language skill at kindergarten entry is related to such skill at first-grade entry. The only other significant correlations are between the SFST oral language tasks and the IRAS reading tasks, which are, interestingly, all negative (ranging from -.28 to -.34), thus suggesting that relatively high Spanish oral language skills at entry to kindergarten are associated with relatively low Spanish reading skills at first-grade entry.

Similar patterns are found for the instruction-intercept correlations (Table 48, bottom panel), but with the expected sign reversal: negative relationships (of comparable magnitude to those in the student-intercepts) between the two sets of oral language tasks, but positive relationships between the SFST oral language tasks and the IRAS reading tasks. For the slope indices (Table 48, top panel), none of the correlations are significant.

The cross language correlations (e.g., Spanish SFST with English IRAS) are presented in Appendix G, and will only be briefly commented on here. For the relationships between the English SFST summary measures and the Spanish IRAS growth indices, only a few of the correlations reach significance. The most notable ones are the negative relationships between the English SFST Vocabulary task and the Spanish IRAS decoding tasks (average value of about -.25). For those relations between the Spanish SFST and the English IRAS measures, the most note-

Table 47

Pre-reading and Reading Measures - Spanish:
Correlations Between IRAS Growth Indices and
SFST Summary Measures for the Kindergarten Bilingual Sample

IRAS - Spanish (N=176)

Slope

		VDC SLOPE	VDF SLOPE	LDC SLOPE	LSP SLOPE	SRD SLOPE	NRC SLOPE	ERC SLOPE	NLC SLOPE	ELC SLOPE
	ALPHPR	0.05	0.06	-0.01	-0.06	0.10	0.10	0.11	0.05	[illegible]
	WRDNAM	0.09	0.10	-0.03	-0.01	0.16	0.16	0.17	0.02	[illegible]
	SNGLTM	-0.21	-0.14	-0.07	-0.18	-0.02	0.01	0.04	[illegible]	[illegible]
SFST	DBLLTM	-0.05	.00	0.12	-0.09	0.01	0.18	0.16	-0.18	[illegible]
Spanish	PSTRNS	0.10	0.06	0.06	0.03	0.16	0.10	0.14	0.08	0.08
Kinder	PSTRNF	0.08	0.10	.00	-0.02	0.17	0.09	0.14	[illegible]	[illegible]
	VCTDCL	0.40	0.17	0.29	0.34	0.48	0.44	0.43	0.31	0.37
	DEFNTN	0.48	0.18	0.40	0.42	0.50	0.40	0.41	0.24	0.32
	CMPPRE	0.36	0.07	0.34	0.29	0.44	0.33	0.36	0.18	0.32

S-Intercept

		VDC S-INT	VDF S-INT	LDC S-INT	LSP S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
	ALPHPR	0.17	0.07	0.19	0.29	0.09	-0.06	-0.05	0.13	-0.12
	WRDNAM	0.18	0.04	0.05	0.16	.00	-0.18	-0.16	0.01	[illegible]
	SNGLTM	0.12	0.11	-0.01	0.16	0.01	-0.07	-0.12	0.09	0.08
SFST	DBLLTM	0.04	0.04	0.10	0.17	0.08	-0.23	-0.24	0.19	[illegible]
Spanish	PSTRNS	0.12	0.11	0.11	0.16	0.06	-0.05	-0.11	0.14	0.12
Kinder	PSTRNF	0.09	0.12	0.08	0.16	0.02	-0.08	[illegible]	0.17	[illegible]
	VCBDCL	0.11	0.33	0.12	0.17	-0.03	-0.32	-0.34	.43	[illegible]
	DEFNTN	0.08	0.27	0.15	0.22	-0.07	-0.28	[illegible]	[illegible]	[illegible]
	CMPPRE	0.16	0.37	0.15	0.27	-0.03	-0.18	-0.29	[illegible]	[illegible]

I-Intercept

		VDC I-INT	VDF I-INT	LDC I-INT	LSP I-INT	SRD I-INT	NRC I-INT	ERC I-INT	NLC I-INT	ELC I-INT
	ALPHPR	-0.16	-0.06	-0.16	-0.09	-0.09	[illegible]	[illegible]	-0.04	[illegible]
	WRDNAM	-0.06	0.11	-0.10	0.07	0.06	0.18	0.15	[illegible]	[illegible]
	SNGLTM	-0.12	-0.02	-0.03	-0.07	0.03	0.01	[illegible]	[illegible]	[illegible]
SFST	DBLLTM	.00	-0.07	-0.04	-0.07	-0.04	0.17	[illegible]	[illegible]	[illegible]
Spanish	PSTRNS	-0.08	-0.06	-0.06	-0.08	.00	0.07	0.11	[illegible]	[illegible]
Kinder	PSTRNF	.00	-0.05	-0.01	-0.11	0.09	0.11	[illegible]	[illegible]	[illegible]
	VCBDCL	-0.01	-0.28	-0.14	-0.04	0.16	0.46	[illegible]	[illegible]	[illegible]
	DEFN	[illegible]	-0.29	-0.13	[illegible]	0.15	0.44	0.34	[illegible]	[illegible]
	CMPPRE	[illegible]	-0.31	-0.10	[illegible]	0.14	0.38	[illegible]	[illegible]	[illegible]

473

BEST COPY AVAILABLE

Table 43

Pre-reading and Reading Measures - Spanish:
Correlations Between IRAS Growth Indices and
SFST Summary Measures for the First-grade Bilingual Sample

IRAS - Spanish (N=60)

Slope

		VDC SLOPE	VDF SLOPE	LDC SLOPE	LSP SLOPE	SRD SLOPE	NRC SLOPE	ERC SLOPE	NLC SLOPE	ELC SLOPE
	ALPHPR	0.25	-0.07	0.13	0.17	0.05	0.09	0.18	-0.06	-0.03
	WRDNAM	-	-	-	-	-	-	-	-	-
	SNGLTM	0.10	-0.14	-0.03	-0.03	[illegible]	-0.01	0.08	-0.08	[illegible]
SFST	DBLLTM	0.27	0.14	0.20	.00	0.04	0.01	0.20	0.13	0.11
Spanish	PSTRNG	0.08	-0.10	-0.11	0.01	0.08	0.06	-0.01	0.09	0.04
First	PSTRNF	0.19	0.22	-0.05	-0.02	0.12	0.14	0.14	0.21	0.20
	VCFDCL	0.27	-0.06	-0.15	0.09	0.03	-0.05	0.11	0.08	0.01
	DEFNTN	-	-	-	-	-	-	-	-	-
	CMPPRB	-	-	-	-	-	-	-	-	-

S-Intercept

		VDC S-INT	VDF S-INT	LDC S-INT	LSP S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
	ALPHPR	0.13	0.27	0.16	0.25	0.13	0.11	0.01	0.29	0.12
	WRDNAM	-	-	-	-	-	-	-	-	-
	SNGLTM	0.17	0.35	0.17	0.23	0.05	0.07	-0.02	0.30	0.14
SFST	DBLLTM	0.16	0.16	0.18	0.34	0.16	0.14	-0.08	0.08	-0.01
Spanish	PSTRNG	0.05	0.23	0.17	0.11	-0.01	.00	0.07	[illegible]	0.10
First	PSTRNF	0.15	0.14	0.28	0.30	0.00	-0.03	-0.04	[illegible]	-0.05
	VCFDCL	0.16	0.51	0.36	0.24	0.13	0.20	0.01	0.29	[illegible]
	DEFNTN	-	-	-	-	-	-	-	-	-
	CMPPRB	-	-	-	-	-	-	-	-	-

I-Intercept

		VDC I-INT	VDF I-INT	LDC I-INT	LSP I-INT	SRD I-INT	NRC I-INT	ERC I-INT	NLC I-INT	ELC I-INT
	ALPHPR	-0.11	-0.16	-0.17	-0.07	.00	-0.01	-0.07	-0.14	-0.17
	WRDNAM	-	-	-	-	-	-	-	-	-
	SNGLTM	-0.17	-0.35	-0.09	-0.17	-0.17	-0.06	0.07	-0.16	[illegible]
SFST	DBLLTM	-0.24	-0.01	-0.30	-0.37	-0.17	-0.09	[illegible]	[illegible]	[illegible]
Spanish	PSTRNG	-0.13	-0.09	-0.12	-0.14	0.22	-0.07	-0.15	-0.21	-0.11
First	PSTRNF	-0.18	-0.07	-0.24	-0.23	.00	-0.02	-0.07	-0.14	[illegible]
	VCFDCL	-0.29	-0.34	-0.28	-0.15	0.02	-0.21	-0.13	-0.21	[illegible]
	DEFNTN	-	-	-	-	-	-	-	-	-
	CMPPRB	-	-	-	-	-	-	-	-	-

474

BEST COPY AVAILABLE

worthy correlations are within the student-intercepts concerning the decoding relationships between the SFST measures of alphabet knowledge, word naming, and letter matching and the IRAS decoding and reading tasks, all of which are positive (average value of about .3). Thus, it seems that relatively high entry skill in decoding, as evidenced in the Spanish SFST, is related to relatively high entry decoding skill on the English IRAS; however, the converse between English and Spanish is not evident.

The correlational pattern between the SFST and IRAS indices can be summarized as follows. First, knowledge of the English alphabet at kindergarten entry is found not only to be generally related to English literacy skill at first-grade entry, but also (positively) to subsequent growth in decoding and reading acquisition. Knowledge of the Spanish alphabet, however, does not carry such widespread predictive power for Spanish literacy development, neither for entry skill nor for subsequent growth. Kindergarten entry skill in decoding and oral language is related to such skills at first-grade entry, within both English and Spanish; but for English, some of these entry skills are further (positively) related to subsequent English literacy growth (segmentation to decoding growth, and oral language to reading growth). For the cross-language correlations, few significant relationships are found.

Finally, the relationships between the SFST and the standardized reading achievement indices will be discussed; the relevant data are presented in Tables 49 and 50 for the English and Spanish measures, respectively. Treating the kindergarten sample, only a few of the relationship for the English SFST are significant, and only that of alphabet knowledge shows a systematic trend of higher alphabet knowledge with higher standardized reading achievement student-intercept and growth. For the Spanish SFST data, relatively high kindergarten skill in alphabet knowledge, word naming, letter matching, phonetic segmentation, definitions, and comprehension are associated with high first-grade entry skill on English standardized reading achievement. This provides further support for the notion that there is some degree of literacy transfer across English and Spanish in this sample of students.

SUMMARY

This section provides a brief summary of the main findings concerning the pre-reading and reading measures discussed in this volume. The summary will follow the general outline of the volume. treating pre-reading, reading, and then the relationship between the two.

Pre-reading Measures

For the pre-reading skills assessed in the English and Spanish versions of the Stanford Foundation Skills Test, the bilingual sample, at kindergarten entry, can be characterized as follows. About half the

Table 49

Pre-reading and Reading Measures - English:
Correlations Between Standardized Reading Achievement Growth Indices
and SFST Summary Measures for the Bilingual Sample

SFST - English Kindergarten
(N=125)

	ALPHPR	WRDNAM	SNGLTM	DBLLTM	PSTRNG	PSTRNF	VCFDCL	DEFNTN	CMPFRB
GEQ SLOPE	0.26	0.17	0.10	.00	0.05	0.13	0.04	-0.04	0.02
GEQ S-INT	0.29	0.20	0.16	0.24	0.18	0.17	0.25	0.25	0.20
GEQ I-INT	-0.07	-0.16	-0.03	-0.12	-0.10	0.04	-0.05	-0.18	-0.16

SFST - English First Grade
(N=48)

	ALPHPR	WRDNAM	SNGLTM	DBLLTM	PSTRNG	PSTRNF	VCFDCL	DEFNTN	CMPFRB
GEQ SLOPE	0.33	-	0.17	0.09	0.25	0.13	0.40	-	-
GEQ S-INT	0.35	-	0.18	0.43	-0.11	0.33	0.38	-	-
GEQ I-INT	-0.08	-	-0.15	-0.18	0.30	-0.17	-0.11	-	-

476

Table E[illegible]

Pre-reading and Reading Measures:
Correlations Between Standardized Reading Achievement - English
and SFST - Spanish Summary Measures for the Bilingual Sample

SFST - Spanish Kindergarten
(N=143)

	ALPHPR	WRDNAM	SNGLTM	DBLLTM	PSTRNG	PSTRNF	VCBDCL	DEFNTN	CMPRHS
SEQ SLOPE	-0.06	-0.1[illegible]	0.03	0.03	0.12	[illegible]	-0.15	-0.24	-0.2[illegible]
SEQ S-INT	0.39	0.41	[illegible]	0.25	0.15	[illegible]	[illegible]	[illegible]	[illegible]
SEQ I-INT	-0.12	.00	-0.10	-0.13	-0.07	-0.11	-0.06	-0.13	[illegible]

SFST - Spanish First Grade
(N=59)

	ALPHPR	WRDNAM	SNGLTM	DBLLTM	PSTRNG	PSTRNF	VCBDCL	DEFNTN	CMPRHS
SEQ SLOPE	0.08	-	0.22	0.24	0.09	0.04	[illegible]	-	-
SEQ S-INT	-0.03	-	0.08	0.25	-0.08	0.29	[illegible]	-	-
SEQ I-INT	0.05	-	0.07	0.03	0.09	-0.19	-0.05	-	-

students come to school knowing the English alphabet, which is found to be a good predictor of early English literacy exposure; knowledge of the Spanish alphabet is negligible, but expected, given its different treatment in the language. Sight word recognition is minimal, but visual matching skills are already highly developed. Auditory segmentation skills can be readily acquired with familiar words by most of the students, but the transfer of this skill to novel items is difficult for some. Vocabulary knowledge is also high at entry, but the formal language dimensions of schooling and text are new to many. In general, these students come to school with sufficient skill to begin literacy acquisition -- they do not seem to be academically disadvantaged.

Reading Achievement Measures

The development of the individual components of reading skill, as assessed in yearly administrations of the English and Spanish Interactive Reading Assessment System, were analyzed as linear growth functions. Such functions describe development by projecting the best-fit line through the yearly obtained data points available for an individual student's performance in a particular task, allowing development to be captured in the parameters of slope and intercept. The slope value provides an estimate of the student's (linear) growth in a component skill area resulting from a single year of instruction, and the intercept provides an estimate of the student's skill level for some fixed entry point in time. Actually, two intercepts were computed in this study: first, the student-intercept, which estimates the level of skill shown by the student at first-grade entry, and the instruction-intercept, which is an estimate of the onset of effective instruction in the component skill area. Coupled with this technique for describing growth, a linear growth track model (fully detailed in Volume 3) was constructed which allows student performance to be compared to expected levels of performance based on the grade-level difficulty of the graded IRAS materials. The following summarizes the performance of the bilingual sample in English reading skills, as expressed by these growth function indices -- the descriptions are based on aggregate performance, and it is important to note that the standard deviations of the these measures indicate substantial individual differences between students.

In the aggregate, the bilingual sample enters first grade with English oral language skills which exceed the expectations of the growth track model, but which grow in accord with the model predictions; thus oral language skills are above grade-level expectations throughout the primary grades. The decoding skills of the sample are minimal at first-grade entry, as expected, and show subsequent growth which is above grade-level expectations (progress in spelling, however, is slow); thus decoding (of isolated words) is, like oral language skill, above expectation throughout the primary grades. Decoding fluency, however, may present problems in reading connected text, as by second-grade exit, the average student still has a reading rate of less than two syllables per second. Reading comprehension is about a half grade level below expectation at entry, and shows growth slightly above the expected rate; thus, reading comprehension is found to be slightly below grade-level expecrations throughout the primary grades. The aggregate picture then, suggests that oral language and decoding skills are being sufficiently developed to allow reading acquisition to proceed in accord with the growth track model; that such is found to be just oelow grade-level expectations may partially be due to fluency in decoding.

Analyses of variance condu on the growth function indices for each of the IRAS scales revealeo gnificant differences based on both site (not summarized here) and on the students' level of English oral skill at kindergarten entry. The differences in performance for the latter appear to be that students who come to school with relatively lower English skills show greater growth in English oral language capacities, and thus, show a convergence in such skill in late fourth grade with those students who entered with higher English skills. However, the high English entry students are better able to profit from decoding instruction in a way that their initial advantage in decoding continues to expand.

Interestingly, the analyses also revealed significant differences in the growth of reading comprehension based on the students' level of Spanish oral skill at kindergarten entry: students' with relatively higher Spanish oral skills at entry to kindergarten have growth rates in English reading comprehension which exceed those of students' with relatively lower entry Spanish oral skills. This suggests that although the development of English listening comprehension does not differ for these groups, relatively higher skills in Spanish at school entry promote the growth of English reading comprehension.

An additional index of English reading growth was provided by the linear growth functions computed over the standardized reading achievement measures (in the form of grade equivalents) collected for each student (such were available for only 80% of the sample). Reading performance, again, in the aggregate, as assessed with these indices indicates that the bilingual sample enters first-grade just slightly below grade-level expectations, and shows growth which is also slightly below expectation -- by fourth grade exit, the sample is projected to be a full grade-level behind.

An analysis of variance conducted on these growth function indices also revealed significant differences based on English oral language entry skill (and also, site -- not summarized here), which can be characterized as follows. The low English entry group begins first-grade just below grade-level expectation, with subsequent growth that gives about three-quarters of a grade-level improvement for each year of instruction. The high English entry group begins first-grade slightly above grade level expectations, and grows at a rate that is slightly below grade-level expectation. Thus, by fourth-grade exit, the high English entry group is projected to be about a half grade level behind, while the low English entry group is more than a full grade level lower.

The performance of the bilingual sample in Spanish reading skills, as revealed in the Spanish IRAS growth functions, can be characterized as follows -- again, this is an aggregate picture, and there are substantial differences for individual students. In the aggregate, the bilingual sample enters first grade with Spanish oral language skills which exceed the expectations of the growth track model, but which grow at half the expected rate; thus oral language skills are above grade-level expectation at entry, but fall below grade level during the primary grades. The decoding skills of the sample are minimal at first-grade entry, as expected, and show subsequent growth which is slightly above grade-level expectations (as in English, however, progress in spelling is slow); thus decoding (of isolated words) is above expectation throughout the primary grades. Also, as in English, the data suggest that decoding fluency may present some difficulties in reading connected text. Reading comprehension is a grade level below expectation at entry, and shows growth which is only half the expected rate; thus, reading comprehension is found to be substantially below grade-level expectation throughout the primary grades. The aggregate picture then, is one where the acquisition of reading comprehension is well below the expectations of the growth track model. Decoding skills for isolated words are close to the growth track, though this skill is not as great as evidenced in the English materials. Oral language skills, which are above expectation at entry to first grade, show growth that is substantially below expectation -- as such, these formal language skills are projected to fall below the growth track, and thus, would seem to provide a major obstacle to Spanish reading acquisition.

Analyses of variance conducted on these growth function indices for each of the IRAS scales revealed significant differences based on the students' level of Spanish oral skill at kindergarten entry, and, as in English, site (again, not summarized here). The Spanish entry differences can be characterized as follows. The low Spanish entry students enter first grade with less skill than the high Spanish group in the areas of formal language and decoding, but subsequent growth does not differ. For reading comprehension, however, the two groups begin with the same low-level skills, but, given the greater formal language and decoding skills of the high Spanish entry group, their growth in reading comprehension is able to proceed at a greater rate. This rate, however, is substantially below that expected from the growth track model, and the data suggest that the major difficulty for these students is not decoding skill, but rather, skill in dealing with the formal language aspects of text.

For both English and Spanish, the relationships found within the reading measures between the component scales can be briefly summarized as follows. First, the highest relationships are generally between the component scales within the three major skill areas assessed (formal language, decoding, and reading). The correlations between these skill areas are strongest for decoding and reading, somewhat weaker between formal language and reading, and weakest between decoding and formal language. Thus, the general correlational pattern suggests that decoding and formal language skills are relatively independent, with both needed for growth in reading comprehension.

The correlational pattern between the English and Spanish indices can be briefly summarized: skills in decoding, and to a lesser degree, in reading, are related across the two languages, while formal language skills (as expected) are generally unrelated. There is, however, a general trend for stronger relationships between a given English task across the set of Spanish tasks when compared to those relationships for the same given Spanish task across the set of English tasks -- this suggests that literacy development in English may be more readily transferable to Spanish than from Spanish to English.

Relationships Between Pre-reading and Reading Achievement Measures

The correlational pattern between the SFST and IRAS indices can be summarized as follows. First, knowledge of the English alphabet at kindergarten entry is found not only to be generally related to English literacy skill at first-grade entry, but also to subsequent growth in decoding and reading acquisition. Knowledge of the Spanish alphabet, however, does not carry such widespread predictive power for Spanish literacy development, neither for entry skill nor for subsequent growth. Kindergarten entry skill in decoding and oral language is related to such skills at first-grade entry, within both English and Spanish; but for English, some of these entry skills are further related to subsequent English literacy growth (segmentation to decoding growth, and oral language to reading growth). For the cross-language correlations, few significant relationships are found.

These accounts of performance have been largely based on the linear growth estimates, and while a substantial amount of growth in the IRAS skills assessed can be explained as linear, there remain, nonetheless, deviations from linearity. Further, the performance descriptions say nothing about the instruction received by the sample. In relating these two dimensions of growth and instruction, the prediction of individual student deviations from the aggregate performance by individual student deviations from the aggregate indices of instruction is of primary import in the study, and is the focus of a subsequent volume. Such analyses should provide considerable insight into the differences in performance described in the volume.

References

Calfee, R. C. (1970). Perceptual and memory components in reading. (Contract No. OEG-9-70-0035). Washington, DC: Department of Health, Education, and Welfare, Office of Education.

Calfee, R. C. (1977). Assessment of independent reading skills: Basic research and practical applications. In A. S. Reber and D. L. Scarborough (Eds.), Toward a psychology of reading. Hillsdale, NJ: Lawrence Erlbaum Associates.

Calfee, R. C., & Associates. (1978). Stanford foundation skills test: Examiner's Manual. Stanford, CA: Stanford University.

Calfee, R. C., & Associates. (1980). Stanford foundation skills test (rev.). Stanford, CA: Stanford University.

Calfee, R. C., & Calfee, K. H. (1979). Interactive reading assessment system. Stanford, CA: Stanford University.

Calfee, R. C., & Calfee, K. H. (1981). Interactive reading assessment system (rev.). Stanford, CA: Stanford University.

Calfee, R. C., Calfee, K. H., & Peña, S. C. (1979). Interactive Reading Assessment System - Spanish. Stanford, CA: Stanford University.

Calfee, R. C., & Drum, P. A. (Eds.). (1979). Teaching reading in compensatory classes. Newark, DE: International Reading Association.

Calfee, R. C., & Peña, S. C. (1978). Prueba Stanford de destrezas fundamentales. Stanford, CA: Stanford University.

Calfee, R. C. & Peña, S. C. (1980). Prueba Stanford de destrezas fundamentales (rev.). Stanford, CA: Stanford University.

Carroll, J. B., Davies, P., & Richman, B. (1971). The American Heritage word frequency book. New York: Houghton Mifflin.

Ehri, L. C. (1979). Linguistic insight: Threshold of reading acquisition. In A. M. Lesgold and C. A. Perfetti (Eds.), Reading research: Advances in theory and practice, Vol. 1. New York, NY: Academic Press.

Gibson, E., & Levin, H. (1975). The psychology of reading. Cambridge, MA: MIT Press.

Gough, P., & Hillinger, M. (1979). Learning to read: An unnatural act. Paper presented at the National Conference of the Orton Society, Indianapolis, Indiana.

Hoover, H. D. (1984). The most appropriate scores for measuring educational development in the elementary schools: GE's. Educational Measurement: Issues and Practice, Winter, 8-14.

Rumelhart, D. (1977). Understanding and summarizing brief stories. In D. Laberge and S. J. Samuels (Eds.), Basic processes in reading: Perception and comprehension. Hillsdale, NJ: Lawrence Erlbaum Associates.

Tunmer, W., Pratt, C., & Herriman, M. (Eds.). (1984). Metalinguistic awareness in children: Theory, research, and implications. Berlin: Springer-Verlag.

Vellutino, F. R. (1980). Dyslexia: Theory and research. Cambridge, MA: MIT Press.

Venezky, R. (1975). The curious role of letter names in reading instruction. Visible Language, 9, 7-23.

158

APPENDICES

APPENDIX A

Informal Reading Inventory/Inventario de la lectura

APPENDIX A

Informal Reading Inventory/Inventario informal de la lectura

INFORMAL READING INVENTORY

An Informal Reading Inventory - IRI was administered to the students in the study on a monthly schedule (revised to each six-week period in the final years of the study) from the time the students began to read connected text. Typically, most first grade children were reading connected text by early January, if not before.

Informal reading inventories were developed by the SEDL research staff in both English and Spanish (Domínguez & Mace-Matluck, 1980; Mace-Matluck & Domínguez, 1978a, 1978b; Mace-Matluck, Domínguez, & Padilla-Hajjar, 1978). These consisted of (a) word lists constructed from words selected randomly from the list of "new words" contained in each of the textbooks included in the instrument and (b) selected passages, of appropriate lengths, taken from the beginning, middle, and end of each of the reading textbooks regularly used in the reading instruction in each of the school districts. The number of words contained in the passages were determined, the required reading time (based on appropriate rates of words per minute) was calculated, and questions were constructed (usually three to five) to assess the students' comprehension of literal facts contained in the passage.

On the basis of the student's ability to read a minimum of one half of the words on the highest ordered list presented, the student was shown the appropriate passages, one at a time, and asked to read the passage aloud. If the student read the passage within the required reading time, the comprehension questions were asked; the student was then asked to read the next passage. This procedure was continued until such time that the student was unable to meet the time criterion. The performance of the student was tape recorded. Subsequently, SEDL staff scored each student's performance for fluency rate (words read per minute), word recognition (percent of words correctly read), and comprehension score (percent of questions answered correctly). A level of reading difficulty was assigned to each passage attempted based on the following criteria:

Level	Word Recognition	Comprehension	Fluency
Independent	97% or more accuracy	80% or more accuracy	Appropriate rate for the material.
Instructional	92% - 96% accuracy	60% - 70% accuracy	Appropriate rate.

Frustrational	91% or lower	50% or lower	Rate below criterion for the material.

References

Domínguez, D., & Mace-Matluck, B. J. (1980). Informal reading inventory/Inventario informal de la lectura. Austin, TX: Southwest Educational Development Laboratory.

Mace-Matluck, B. J., & Domínguez, D. (1978a). Informal reading inventory: Keys to reading. Austin, TX: Southwest Educational Development Laboratory.

Mace-Matluck, B. J., & Domínguez, D. (1978b). Inventario informal de la lectura. Austin, TX: Southwest Educational Development Laboratory.

Mace-Matluck, B. J., Domínguez, D., & Padilla-Hajjar, R. (1978). Inventario informal de la lectura. Austin, TX: Southwest Educational Development Laboratory.

458

APPENDIX B

Interactive Reading Assessment System - English:

Descriptive Statistics on Growth Indices for the Bilingual Sample for Each Site Overall and by Language Category

Table 1

Interactive Reading Assessment System - English:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Non-comprehension Scales for Site 0 Overall by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-0	Lo Hi-0	Hi Lo-0	Hi Hi-0
VDCE	Slope	M	2.7	1.0	2.6		3.2
VDCE	Slope	S	1.9	1.0	1.6		2.0
VDCE	Slope	N	51	9	13		29
VDCE	S-Intrcp	M	1.3	-0.5	0.8		2.1
VDCE	S-Intrcp	S	2.7	1.8	2.7		2.6
VDCE	I-Intrcp	M	0.5	[illegible]	0.5		0.2
VDCE	I-Intrcp	S	1.2	1.0	1.1		1.2
VDFE	Slope	M	1.9	1.1	2.1		2.0
VDFE	Slope	S	1.8	1.3	2.0		1.8
VDFE	Slope	N	50	9	13		28
VDFE	S-Intrcp	M	3.8	1.8	3.4		4.6
VDFE	S-Intrcp	S	3.3	3.3	3.9		2.8
VDFE	I-Intrcp	M	-1.4	-0.3	-1.8		-1.7
VDFE	I-Intrcp	S	2.4	2.5	2.8		2.2
LDCE	Slope	M	1.5	1.2	1.8		1.5
LDCE	Slope	S	1.1	1.8	1.0		0.9
LDCE	Slope	N	51	9	13		29
LDCE	S-Intrcp	M	0.3	-1.6	-0.4		1.1
LDCE	S-Intrcp	S	3.2	3.3	1.1		3.6
LDCE	I-Intrcp	M	0.1	1.2	1.0		-0.6
LDCE	I-Intrcp	S	2.4	1.3	0.6		2.9
LSPE	Slope	M	11.8	8.1	13.1		12.3
LSPE	Slope	S	8.6	7.2	8.1		9.1
LSPE	Slope	N	[illegible]	[illegible]	13		29
LSPE	S-Intrcp	M	[illegible]	[illegible]	-0.9		6.7
LSPE	S-Intrcp	S	20.2	11.9	16.4		22.6
LSPE	I-Intrcp	M	0.5	1.5	0.6		0.1
LSPE	I-Intrcp	S	1.9	1.2	1.7		2.1
SRDE	Slope	M	0.5	0.4	0.6		0.6
SRDE	Slope	S	0.4	0.3	0.2		0.5
SRDE	Slope	N	49	8	13		28
SRDE	S-Intrcp	M	0.5	-0.6	0.2		0.9
SRDE	S-Intrcp	S	1.4	0.8	0.8		1.5
SRDE	I-Intrcp	M	-0.1	2.0	0.1		-0.7
SRDE	I-Intrcp	S	2.4	1.7	[illegible]		2.4

430

Table 2

Interactive Reading Assessment System - English:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Comprehension Scales for Site 0 Overall and by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-0	Lo Hi-0	Hi Lo-0	Hi Hi-0
NRCE	Slope	M	1.6	0.6	1.8		1.7
NRCE	Slope	S	1.1	1.3	0.7		1.1
NRCE	Slope	N	51	9	13		29
NRCE	S-Intrcp	M	-0.3	-1.5	-1.3		0.5
NRCE	S-Intrcp	S	2.5	3.6	1.8		2.2
NRCE	I-Intrcp	M	0.8	0.7	1.8		0.3
NRCE	I-Intrcp	S	1.3	1.5	1.1		1.0
ERCE	Slope	M	1.6	0.8	2.0		1.7
ERCE	Slope	S	1.7	1.5	1.6		1.7
ERCE	Slope	N	49	8	13		28
ERCE	S-Intrcp	M	-0.4	-2.2	-2.0		0.8
ERCE	S-Intrcp	S	4.8	4.1	3.8		5.2
ERCE	I-Intrcp	M	0.1	0.9	1.1		-0.6
ERCE	I-Intrcp	S	2.7	1.7	2.1		3.0
NLCE	Slope	M	1.5	1.2	1.7		1.4
NLCE	Slope	S	0.9	1.2	0.8		0.8
NLCE	Slope	N	51	9	13		29
NLCE	S-Intrcp	M	1.9	0.3	0.7		2.8
NLCE	S-Intrcp	S	2.2	1.1	2.1		1.9
NLCE	I-Intrcp	M	-0.6	0.9	0.2		-1.4
NLCE	I-Intrcp	S	2.0	1.1	1.9		2.0
ELCE	Slope	M	1.1	0.5	1.9		1.0
ELCE	Slope	S	1.3	1.7	1.2		1.1
ELCE	Slope	N	49	8	13		28
ELCE	S-Intrcp	M	2.5	1.2	.0		3.9
ELCE	S-Intrcp	S	4.1	5.3	3.8		3.2
ELCE	I-Intrcp	M	-0.8	1.2	0.4		-1.9
ELCE	I-Intrcp	S	2.8	1.6	2.6		2.7

Table 3

Interactive Reading Assessment System - English:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Non-comprehension Scales for Site 1 Overall by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-1	Lo Hi-1	Hi Lo-1	Hi Hi-1
VDCE	Slope	M	4.4	3.3	3.9	7.0	3.4
VDCE	Slope	S	2.9	0.9	1.2	5.1	1.9
VDCE	Slope	N	17	4	6	4	3
VDCE	S-Intrcp	M	-1.7	-2.0	-2.0	-2.3	0.2
VDCE	S-Intrcp	S	3.0	4.1	1.5	4.5	2.0
VDCE	I-Intrcp	M	1.2	1.4	1.5	1.1	0.7
VDCE	I-Intrcp	S	0.9	1.2	0.2	0.8	1.0
VDFE	Slope	M	1.8	1.6	2.3	1.7	0.9
VDFE	Slope	S	1.3	1.4	1.3	1.3	1.2
VDFE	Slope	N	17	4	6	4	3
VDFE	S-Intrcp	M	3.9	2.6	2.0	6.2	6.0
VDFE	S-Intrcp	S	3.1	3.8	2.1	1.4	3.2
VDFE	I-Intrcp	M	-1.3	-2.1	-0.9	-1.0	-1.4
VDFE	I-Intrcp	S	2.2	3.2	2.6	0.7	2.0
LDCE	Slope	M	1.5	2.1	1.9	0.9	1.0
LDCE	Slope	S	0.9	0.3	0.9	1.1	0.2
LDCE	Slope	N	17	4	6	4	3
LDCE	S-Intrcp	M	1.4	-1.1	0.9	3.4	3.3
LDCE	S-Intrcp	S	2.5	1.3	1.6	3.2	1.0
LDCE	I-Intrcp	M	-0.4	1.5	-0.2	-0.9	-2.7
LDCE	I-Intrcp	S	2.2	0.6	2.4	2.0	1.7
LSPE	Slope	M	13.6	13.3	8.6	19.9	16.5
LSPE	Slope	S	11.7	5.1	5.8	22.6	5.2
LSPE	Slope	N	17	4	6	4	3
LSPE	S-Intrcp	M	4.3	-5.3	-0.1	15.7	10.8
LSPE	S-Intrcp	S	21.8	20.6	7.9	40.0	5.3
LSPE	I-Intrcp	M	0.2	1.3	0.6	-1.6	0.3
LSPE	I-Intrcp	S	2.3	1.6	2.0	3.6	0.4
SRDE	Slope	M	0.7	0.5	0.7	0.9	0.6
SRDE	Slope	S	0.7	0.6	0.2	0.4	0.1
SRDE	Slope	N	17	4	6	4	3
SRDE	S-Intrcp	M	-0.1	-0.1	-0.5	0.2	0.1
SRDE	S-Intrcp	S	1.0	2.1	0.5	0.6	0.4
SRDE	I-Intrcp	M	1.2	1.5	1.6	0.5	0.8
SRDE	I-Intrcp	S	1.0	1.4	0.6	1.1	0.5

492

Table 4

Interactive Reading Assessment System - English:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Comprehension Scales for Site 1 Overall and by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-1	Lo Hi-1	Hi Lo-1	Hi Hi-1
NRCE	Slope	M	3.4	4.1	3.7	2.4	3.3
NRCE	Slope	S	1.7	2.1	1.5	1.2	2.2
NRCE	Slope	N	17	4	6	4	3
NRCE	S-Intrcp	M	-5.2	-8.8	-6.4	-0.6	-3.9
NRCE	S-Intrcp	S	5.9	8.5	5.2	1.7	5.3
NRCE	I-Intrcp	M	2.1	2.7	2.6	1.2	1.8
NRCE	I-Intrcp	S	1.0	1.2	0.7	0.6	0.8
ERCE	Slope	M	3.0	3.9	3.2	2.5	2.3
ERCE	Slope	S	1.8	2.8	1.9	0.4	1.4
ERCE	Slope	N	17	4	6	4	3
ERCE	S-Intrcp	M	-5.1	-8.7	-6.0	-2.3	-2.0
ERCE	S-Intrcp	S	6.0	9.8	5.1	0.8	4.1
ERCE	I-Intrcp	M	2.2	2.5	2.7	2.0	1.2
ERCE	I-Intrcp	S	1.2	1.8	0.7	0.4	1.8
NLCE	Slope	M	1.3	1.6	1.2	1.4	0.8
NLCE	Slope	S	0.6	0.7	0.7	0.1	0.5
NLCE	Slope	N	17	4	6	4	3
NLCE	S-Intrcp	M	2.3	0.6	2.4	2.5	4.0
NLCE	S-Intrcp	S	2.5	3.4	2.5	1.3	1.9
NLCE	I-Intrcp	M	-1.5	-0.3	-1.8	-0.8	-3.5
NLCE	I-Intrcp	S	2.5	2.7	2.9	0.9	2.7
ELCE	Slope	M	1.8	2.3	2.0	1.4	1.4
ELCE	Slope	S	0.7	1.2	0.4	0.2	0.1
ELCE	Slope	N	17	4	6	4	3
ELCE	S-Intrcp	M	0.4	-1.6	-0.2	2.0	1.9
ELCE	S-Intrcp	S	2.9	4.8	2.3	1.2	0.5
ELCE	I-Intrcp	M	0.3	0.7	0.9	-0.6	-0.4
ELCE	I-Intrcp	S	1.5	2.5	1.1	1.2	0.4

Table 5

Interactive Reading Assessment System - English:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Non-comprehension Scales for Site 2 Overall by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-2	Lo Hi-2	Hi Lo-2	Hi Hi-2
VDCE	Slope	M	2.7	2.0	2.5	3.3	2.8
VDCE	Slope	S	1.2	1.0	1.2	1.1	1.4
VDCE	Slope	N	36	3	16	4	13
VDCE	S-Intrcp	M	-0.1	-0.4	-1.7	1.2	1.7
VDCE	S-Intrcp	S	2.6	1.9	1.8	3.3	2.0
VDCE	I-Intrcp	M	1.0	1.3	1.8	0.4	0.1
VDCE	I-Intrcp	S	1.2	1.1	0.7	1.2	1.0
VDFE	Slope	M	1.5	1.7	2.2	0.8	0.9
VDFE	Slope	S	1.1	1.7	0.8	0.8	0.9
VDFE	Slope	N	36	3	16	4	13
VDFE	S-Intrcp	M	2.8	1.5	-0.6	6.3	6.1
VDFE	S-Intrcp	S	4.1	7.0	2.3	2.5	1.3
VDFE	I-Intrcp	M	-0.6	0.9	1.1	-2.6	-2.5
VDFE	I-Intrcp	S	2.5	1.6	1.2	2.8	2.0
LDCE	Slope	M	2.1	0.7	1.7	3.4	2.6
LDCE	Slope	S	1.3	1.1	1.2	1.8	0.9
LDCE	Slope	N	36	3	16	4	13
LDCE	S-Intrcp	M	-1.3	0.4	-1.5	-1.3	-1.3
LDCE	S-Intrcp	S	2.0	0.7	1.8	2.1	2.3
LDCE	I-Intrcp	M	1.2	-1.3	1.6	1.3	1.3
LDCE	I-Intrcp	S	1.5	3.2	1.3	0.6	0.7
LSPE	Slope	M	10.4	16.2	6.6	9.5	14.2
LSPE	Slope	S	7.7	7.6	6.4	7.2	7.3
LSPE	Slope	N	36	3	16	4	13
LSPE	S-Intrcp	M	-5.6	-23.7	-7.6	10.5	-3.9
LSPE	S-Intrcp	S	15.0	15.9	14.6	14.8	11.0
LSPE	I-Intrcp	M	1.2	2.7	1.8	-0.7	0.8
LSPE	I-Intrcp	S	2.1	1.1	2.0	2.9	1.9
SRDE	Slope	M	0.8	0.6	0.9	0.7	0.7
SRDE	Slope	S	0.4	0.5	0.4	0.5	0.4
SRDE	Slope	N	35	3	15	4	13
SRDE	S-Intrcp	M	0.2	0.3	-0.6	0.7	0.9
SRDE	S-Intrcp	S	1.1	0.4	0.9	0.9	0.9
SRDE	I-Intrcp	M	0.4	-1.2	1.7	-0.6	-0.4
SRDE	I-Intrcp	S	1.9	3.3	0.9	1.5	1.6

494

Table 6

Interactive Reading Assessment System - English:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Comprehension Scales for Site 2 Overall and by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-2	Lo Hi-2	Hi Lo-2	Hi Hi-2
NRCE	Slope	M	2.0	1.7	2.2	1.	1.9
NRCE	Slope	S	1.0	0.5	1.2	1.4	0.5
NRCE	Slope	N	36	3	16	4	13
NRCE	S-Intrcp	M	-0.9	-1.1	-2.6	1.0	0.7
NRCE	S-Intrcp	S	2.7	1.4	2.7	2.9	1.5
NRCE	I-Intrcp	M	1.1	1.8	1.9	-0.1	0.4
NRCE	I-Intrcp	S	1.4	1.2	1.1	1.6	1.2
ERCE	Slope	M	2.4	1.6	2.9	1.2	2.3
ERCE	Slope	S	1.3	1.2	1.5	1.3	0.8
ERCE	Slope	N	35	3	15	4	13
ERCE	S-Intrcp	M	-1.9	-1.7	-4.6	2.5	-0.2
ERCE	S-Intrcp	S	4.4	1.5	4.7	5.1	2.1
ERCE	I-Intrcp	M	1.3	2.5	2.3	-1.4	0.7
ERCE	I-Intrcp	S	2.1	1.1	1.0	4.2	1.4
NLCE	Slope	M	1.7	1.8	2.0	2.1	1.2
NLCE	Slope	S	0.9	0.6	0.8	1.8	0.7
NLCE	Slope	N	36	3	16	4	13
NLCE	S-Intrcp	M	0.9	-0.5	-0.9	2.0	3.1
NLCE	S-Intrcp	S	2.8	3.3	1.6	3.4	2.0
NLCE	I-Intrcp	M	-0.1	0.9	1.5	-1.5	-1.7
NLCE	I-Intrcp	S	2.3	2.0	0.9	2.7	2.2
ELCE	Slope	M	2.1	2.1	2.7	1.8	1.4
ELCE	Slope	S	1.3	0.4	1.2	2.2	0.9
ELCE	Slope	N	35	3	15	4	13
ELCE	S-Intrcp	M	-0.2	-2.0	-3.1	2.9	2.6
ELCE	S-Intrcp	S	4.1	1.3	3.3	4.9	2.5
ELCE	I-Intrcp	M	0.1	2.0	1.9	-2.1	-1.9
ELCE	I-Intrcp	S	2.7	0.8	1.4	3.4	2.2

Table 7

Interactive Reading Assessment System - English:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Non-comprehension Scales for Site 3 Overall by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-3	Lo Hi-3	Hi Lo-3	Hi Hi-3
VDCE	Slope	M	3.7	3.2		3.8	4.9
VDCE	Slope	S	2.0	1.7		1.8	3.1
VDCE	Slope	N	73	28		36	9
VDCE	S-Intrcp	M	-1.1	-1.5		-0.7	-1.7
VDCE	S-Intrcp	S	2.5	1.6		2.7	3.3
VDCE	I-Intrcp	M	1.0	1.2		0.8	1.3
VDCE	I-Intrcp	S	1.4	1.1		1.7	0.5
VDFE	Slope	M	1.8	2.1		1.3	2.7
VDFE	Slope	S	2.5	1.9		2.5	3.6
VDFE	Slope	N	73	28		36	9
VDFE	S-Intrcp	M	4.1	3.1		5.0	3.1
VDFE	S-Intrcp	S	3.6	3.0		3.4	5.0
VDFE	I-Intrcp	M	-1.4	-1.5		-1.5	-0.6
VDFE	I-Intrcp	S	2.6	2.9		2.6	2.5
LDCE	Slope	M	2.5	2.4		2.5	3.2
LDCE	Slope	S	2.0	1.9		2.0	2.3
LDCE	Slope	N	73	28		36	9
LDCE	S-Intrcp	M	-1.0	-1.5		-0.6	-1.3
LDCE	S-Intrcp	S	2.3	1.5		2.6	3.3
LDCE	I-Intrcp	M	1.0	1.3		0.8	0.6
LDCE	I-Intrcp	S	1.4	0.8		1.4	2.2
LSPE	Slope	M	18.2	12.2		20.9	25.7
LSPE	Slope	S	14.7	9.6		14.7	20.9
LSPE	Slope	N	67	26		32	9
LSPE	S-Intrcp	M	-8.7	-7.3		-9.1	-11.1
LSPE	S-Intrcp	S	12.7	9.4		13.1	19.6
LSPE	I-Intrcp	M	1.3	1.3		1.3	1.1
LSPE	I-Intrcp	S	0.7	0.7		0.7	0.7
SRDE	Slope	M	1.3	1.2		1.2	1.4
SRDE	Slope	S	0.7	0.7		0.7	0.7
SRDE	Slope	N	73	28		36	9
SRDE	S-Intrcp	M	-0.5	-0.7		-0.3	-0.6
SRDE	S-Intrcp	S	0.9	0.6		1.1	1.1
SRDE	I-Intrcp	M	1.1	1.4		0.9	1.1
SRDE	I-Intrcp	S	1.2	0.6		1.5	0.9

Table 8

Interactive Reading Assessment System - English:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Comprehension Scales for Site 3 Overall and by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-3	Lo Hi-3	Hi Lo-3	Hi Hi-3
NRCE	Slope	M	2.4	2.3		2.4	3.0
NRCE	Slope	S	1.9	1.9		1.8	2.3
NRCE	Slope	N	73	28		36	9
NRCE	S-Intrcp	M	-1.0	-1.		-0.6	-1.2
NRCE	S-Intrcp	S	2.2	1		2.4	3.2
NRCE	I-Intrcp	M	1.0	1		0.9	0.9
NRCE	I-Intrcp	S	1.1	0.8		1.3	1.4
ERCE	Slope	M	2.2	2.1		2.3	2.1
ERCE	Slope	S	1.8	1.8		1.8	2.1
ERCE	Slope	N	73	28		36	9
ERCE	S-Intrcp	M	-1.0	-1.5		-0.7	-0.6
ERCE	S-Intrcp	S	2.0	1.3		2.2	3.0
ERCE	I-Intrcp	M	0.9	1.0		0.8	0.9
ERCE	I-Intrcp	S	1.3	0.8		1.6	1.0
NLCE	Slope	M	1.7	2.4		1.4	1.0
NLCE	Slope	S	1.9	2.2		1.7	1.4
NLCE	Slope	N	72	28		35	9
NLCE	S-Intrcp	M	2.6	0.7		3.7	4.1
NLCE	S-Intrcp	S	3.3	3.2		2.8	2.5
NLCE	I-Intrcp	M	-1.4	.0		-2.2	-2.4
NLCE	I-Intrcp	S	2.6	2.1		2.5	2.6
ELCE	Slope	M	2.2	2.7		1.8	1.9
ELCE	Slope	S	2.0	2.1		2.0	1.8
ELCE	Slope	N	73	28		36	9
ELCE	S-Intrcp	M	1.1	-0.3		2.0	2.3
ELCE	S-Intrcp	S	3.2	3.1		2.9	3.6
ELCE	I-Intrcp	M	-0.3	0.4		-0.6	-1.5
ELCE	I-Intrcp	S	2.1	1.5		2.1	2.8

Table 9

Interactive Reading Assessment System - English:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Non-comprehension Scales for Site 5 Overall by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-5	Lo Hi-5	Hi Lo-5	Hi Hi-5
VDCE	Slope	M	[illegible].0	2.0	3.7	3.5	3.0
VDCE	Slope	S	2.2	2.0	2.7	2.3	1.7
VDCE	Slope	N	69	15	19	9	27
VDCE	S-Intrcp	M	0.3	0.3	-1 1	1.1	1.1
VDCE	S-Intrcp	S	3.0	2.9	3.1	3.1	2.9
VDCE	I-Intrcp	M	0.4	0.1	1.1	0.1	0.1
VDCE	I-Intrcp	S	1.7	2.2	0.7	2.2	1.8
VDFE	Slope	M	2.1	1.9	3.5	1.8	1.4
VDFE	Slope	S	2.5	2.6	2.4	2.7	2.1
VDFE	Slope	N	67	15	17	8	27
VDFE	S-Intrcp	M	3.2	2.4	0.9	4.9	4.6
VDFE	S-Intrcp	S	4.2	4.3	4.2	3.2	5.7
VDFE	I-Intrcp	M	-1.1	-0.4	-0.5	-1.0	-2.0
VDFE	I-Intrcp	S	2.7	2.5	2.6	2.5	2.8
LDCE	Slope	M	1.8	1.6	2.8	0.9	1.4
LDCE	Slope	S	2.2	2.1	2.7	1.4	1.9
LDCE	Slope	N	65	15	19	7	24
LDCE	S-Intrcp	M	1.1	0.2	-0.4	3.4	2.1
LDCE	S-Intrcp	S	3.7	2.8	4.1	3.7	3.5
LDCE	I-Intrcp	M	0.2	0.7	0.7	-1.2	0.0
LDCE	I-Intrcp	S	1.8	0.8	1.6	2 8	1.9
LSPE	Slope	M	11.3	6.5	10.2	7.8	15.8
LSPE	Slope	S	13.0	8.4	10.9	8.3	16.3
LSPE	Slope	N	69	15	19	8	27
LSPE	S-Intrcp	M	6.9	5.5	1.8	16.7	7.8
LSPE	S-Intrcp	S	19.8	17.1	16.6	20.1	22.[illegible]
LSPE	I-Intrcp	M	0.5	0.7	0.5	-0.1	0.5
LSPE	I-Intrcp	S	1.3	0.8	1.6	1.1	1.3
SRDE	Slope	M	0.9	0.4	1.1	0.7	1.0
SRDE	Slope	S	0.7	0.5	0.8	0.3	0.6
SRDE	Slope	N	69	15	19	8	27
SRDE	S-Intrcp	M	-0.1	0.1	-0.6	0.3	0.0
SRDE	S-Intrcp	S	0.7	0.6	0.6	0.5	0.7
SRDE	I-Intrcp	M	0.7	0.7	1.[illegible]	0.3	0.5
SRDE	I-Intrcp	S	1.3	0.7	0.7	1.6	1.7

498

Table 10

Interactive Reading Assessment System - English:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Comprehension Scales for Site 5 Overall and by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-5	Lo Hi-5	Hi Lo-5	Hi Hi-5
NRCE	Slope	M	2.0	0.9	2.5	2.0	2.3
NRCE	Slope	S	1.8	1.3	2.2	1.4	1.7
NRCE	Slope	N	69	15	19	8	27
NRCE	S-Intrcp	M	-0.6	.0	-1.6	0.0	-0.4
NRCE	S-Intrcp	S	2.0	1.1	1.8	2.0	2.3
NRCE	I-Intrcp	M	0.6	0.2	1.0	0.9	0.4
NRCE	I-Intrcp	S	1.3	0.8	1.1	0.8	1.8
ERCE	Slope	M	2.0	0.9	2.2	2.2	2.5
ERCE	Slope	S	2.0	1.5	2.2	1.6	1.9
ERCE	Slope	N	69	15	19	8	27
ERCE	S-Intrcp	M	-0.8	-0.1	-1.4	-0.5	-0.8
ERCE	S-Intrcp	S	1.9	1.1	1.9	1.7	2.2
ERCE	I-Intrcp	M	0.6	0.2	1.0	0.7	0.6
ERCE	I-Intrcp	S	1.3	0.8	1.1	0.9	1.6
NLCE	Slope	M	1.8	1.6	2.4	1.5	1.6
NLCE	Slope	S	1.6	1.5	1.6	1.6	1.5
NLCE	Slope	N	69	15	19	8	27
NLCE	S-Intrcp	M	2.0	1.5	0.7	2.7	3.0
NLCE	S-Intrcp	S	3.0	2.5	3.1	2.8	3.1
NLCE	I-Intrcp	M	-0.8	0.3	-0.1	-0.4	-2.1
NLCE	I-Intrcp	S	2.4	1.2	2.2	2.0	2.7
ELCE	Slope	M	2.1	1.6	3.3	1.9	1.6
ELCE	Slope	S	1.9	2.1	2.1	1.1	1.4
ELCE	Slope	N	69	15	19	8	27
ELCE	S-Intrcp	M	1.0	0.4	-1.2	2.4	2.4
ELCE	S-Intrcp	S	3.1	[illegible]	2.8	2.3	2.9
ELCE	I-Intrcp	M	[illegible]	0.2	0.8	-0.2	-0.6
ELCE	I-Intrcp	S	1.9	1.9	1.8	1.3	2.1

APPENDIX C

Interactive Reading Assessment System - English:

Summary Tables for the Analyses of Variance on Growth Indices for the Bilingual Sample

Table 1

Interactive Reading Assessment System - English

ANOVA Summary Table for Real Word Decoding Growth Indices
for the Bilingual Sample

Growth Index	Variable	Sum of Squares	df	Mean Square	F	p
Slope	English	20.44	1	20.44	5.46	0.020
	Spanish	13.95	1	13.95	3.72	0.055
	Site	81.80	4	20.45	5.46	0.[illegible]
	Eng x Spn	30.92	1	30.92	8.26	0.004
	Eng x Site	17.38	4	4.35	1.16	0.329
	Error	835.24	223	3.75		
E-intrcp	English	147.26	1	147.26	21.32	0.001
	Spanish	3.10	1	3.10	0.45	0.503
	Site	206.54	4	51.63	7.47	0.001
	Eng x Spn	13.75	1	13.75	1.99	0.160
	Eng x Site	16.96	4	4.24	0.61	0.652
	Error	1540.62	223	6.91		
I-intrcp	English	29.25	1	29.25	15.28	0.001
	Spanish	2.28	1	2.28	1.23	0.268
	Site	26.96	4	6.74	3.65	0.007
	Eng x Spn	3.75	1	3.75	2.03	0.156
	Eng x Site	9.34	4	2.34	1.19	0.242
	Error	412.21	223	1.85		

Table 2

Interactive Reading Assessment System - English:

ANOVA Summary Table for Vocabulary Definition Growth Indices for the Bilingual Sample

Growth Index	Variable	Sum of Squares	df	Mean Square	F	p
Slope	English	38.70	1	38.70	8.92	0.003
	Spanish	24.81	1	24.81	5.72	0.018
	Site	18.14	4	4.54	1.05	0.385
	Eng x Spn	19.89	1	19.88	4.58	0.033
	Eng x Site	25.11	4	6.28	1.45	0.220
	Error	954.66	220	4.34		
S-intrcp	English	547.98	1	547.98	46.79	0.001
	Spanish	40.20	1	40.20	3.43	0.065
	Site	25.03	4	6.26	0.53	0.711
	Eng x Spn	7.80	1	7.80	0.67	0.415
	Eng x Site	123.31	4	30.83	2.63	0.035
	Error	2577.29	220	11.72		
I-intrcp	English	53.71	1	53.71	8.70	0.004
	Spanish	0.01	1	0.01	.00	0.969
	Site	13.23	4	3.31	0.54	0.713
	Eng x Spn	3.12	1	3.12	0.51	0.479
	Eng x Site	81.19	4	20.30	3.29	0.012
	Error	1358.50	220	6.18		

Table 3

Interactive Reading Assessment System (English)

ANOVA Summary Table for Letter-Sound Decoding Growth Indices
for the Bilingual Sample

Growth Index	Variable	Sum of Squares	df	Mean Square	F	p
Slope	English	1.90	1	1.90	0.63	0.430
	Spanish	8.91	1	8.91	2.94	0.088
	Site	49.59	4	12.40	4.09	0.003
	Eng x Spn	5.03	1	5.03	1.66	0.199
	Eng x Site	33.68	4	8.42	2.78	0.028
	Error	664.45	219	3.03		
S-intrcp	English	140.74	1	140.74	17.94	0.001
	Spanish	1.37	1	1.37	0.18	0.676
	Site	263.01	4	65.73	8.38	0.001
	Eng x Spn	0.23	1	0.23	0.03	0.865
	Eng x Site	45.09	4	11.27	1.44	0.223
	Error	1718.53	219	7.65		
I-intrcp	English	46.42	1	46.42	15.24	0.001
	Spanish	0.08	1	0.08	0.03	0.872
	Site	62.99	4	15.75	5.17	0.001
	Eng x Spn	0.31	1	0.31	0.10	0.750
	Eng x Site	16.48	4	4.12	1.35	0.252
	Error	667.22	219	3.05		

503

BEST COPY AVAILABLE

Table 4

Interactive Reading Assessment System - English:

Summary Table for Letter-Sound Spelling Growth Indices
for the Bilingual Sample

Growth Index	Variable	Sum of Squares	df	Mean Square	F	P
Slope	English	1598.12	1	1598.12	11.64	0.001
	Spanish	269.85	1	269.85	1.97	0.162
	Site	2068.92	4	517.23	3.77	0.006
	Eng x Spn	216.24	1	216.24	1.58	0.211
	Eng x Site	691.14	4	172.78	1.26	0.287
	Error	29787.27	217	137.27		
S-intrcp	English	2316.43	1	2316.43	7.77	0.006
	Spanish	210.96	1	210.96	0.71	0.401
	Site	9810.90	4	2452.73	8.23	0.001
	Eng x Spn	899.26	1	899.26	3.02	0.084
	Eng x Site	2164.29	4	541.07	1.82	0.127
	Error	64679.68	217	298.06		
I-intrcp	English	19.72	1	19.72	8.59	0.004
	Spanish	0.40	1	0.40	0.17	0.677
	Site	41.47	4	10.37	4.52	0.002
	Eng x Spn	12.10	1	12.10	5.27	0.023
	Eng x Site	25.31	4	6.33	2.76	0.029
	Error	498.13	217	2.30		

Table 5

Interactive Reading Assessment System - English:

ANOVA Summary Table for Sentence Reading Growth Indices
for the Bilingual Sample

Growth Index	Variable	Sum of Squares	df	Mean Square	F	P
Slope	English	0.01	1	0.01	0.03	0.870
	Spanish	3.09	1	3.09	8.86	0.003
	Site	19.51	4	4.88	13.97	0.001
	Eng x Spn	1.14	1	1.14	3.25	0.073
	Eng x Site	0.46	4	0.12	0.33	0.859
	Error	76.81	220	0.35		
E-intrcp	English	27.43	1	27.43	30.96	0.001
	Spanish	1.85	1	1.85	2.09	0.150
	Site	29.47	4	7.37	8.32	0.001
	Eng x Spn	0.98	1	0.98	1.11	0.293
	Eng x Site	5.17	4	1.29	1.46	0.216
	Error	194.92	220	0.89		
I-intrcp	English	46.44	1	46.44	20.30	0.001
	Spanish	1.28	1	1.28	0.56	0.455
	Site	42.49	4	10.62	4.64	0.001
	Eng x Spn	1.72	1	1.72	0.75	0.387
	Eng x Site	8.51	4	2.13	0.93	0.448
	Error	503.40	220	2.29		

595

Table 5

Interactive Reading Assessment System - English:

ANOVA Summary Table for Narrative Reading Comprehension Growth Indices for the Bilingual Sample

Growth Index	Variable	Sum of Squares	df	Mean Square	F	p
Slope	English	0.01	1	0.01	.00	0.954
	Spanish	27.81	1	27.81	10.64	0.001
	Site	79.08	4	19.77	7.57	0.001
	Eng x Spn	3.74	1	3.74	1.43	0.233
	Eng x Site	7.54	4	1.88	0.72	0.578
	Error	582.70	223	2.61		
S-intrcp	English	174.97	1	174.97	26.71	0.001
	Spanish	28.81	1	28.81	4.40	0.037
	Site	334.96	4	83.74	12.78	0.001
	Eng x Spn	0.20	1	0.20	0.03	0.862
	Eng x Site	79.15	4	19.79	3.02	0.019
	Error	1461.01	223	6.55		
I-intrcp	English	27.44	1	27.44	19.20	0.001
	Spanish	4.99	1	4.99	3.49	0.063
	Site	37.02	4	9.25	6.48	0.001
	Eng x Spn	3.52	1	3.52	2.47	0.118
	Eng x Site	9.63	4	2.41	1.69	0.154
	Error	318.62	223	1.43		

Table 7

Interactive Reading Assessment System - English:

ANOVA Summary Table for Expository Reading Comprehension Growth Indices for the Bilingual Sample

Growth Index	Variable	Sum of Squares	df	Mean Square	F	p
Slope	English	0.05	1	0.05	0.02	0.904
	Spanish	18.11	1	18.11	5.92	0.016
	Site	42.06	4	10.52	3.44	0.009
	Eng x Spn	3.67	1	3.67	1.20	0.274
	Eng x Site	10.41	4	2.60	0.85	0.494
	Error	672.33	220	3.06		
S-intrcp	English	210.10	1	210.10	19.52	0.001
	Spanish	22.21	1	22.21	2.06	0.152
	Site	288.15	4	72.04	6.69	0.001
	Eng x Spn	3.62	1	3.62	0.34	0.562
	Eng x Site	101.00	4	25.25	2.35	0.056
	Error	2367.51	220	10.76		
I-intrcp	English	40.29	1	40.29	13.92	0.001
	Spanish	5.71	1	5.71	1.97	0.162
	Site	71.49	4	17.87	6.18	0.001
	Eng x Spn	1.52	1	1.52	0.52	0.470
	Eng x Site	20.06	4	5.02	1.73	0.144
	Error	636.58	220	2.89		

507

Table 8

Interactive Reading Assessment System - English:

ANOVA Summary Table for Narrative Listening Comprehension Growth Indices
for the Bilingual Sample

Growth index	Variable	Sum of Squares	df	Mean Square	F	P
Slope	English	20.25	1	20.25	9.89	0.002
	Spanish	0.53	1	0.53	0.26	0.610
	Site	8.64	4	2.16	1.06	0.380
	Eng x Spn	4.59	1	4.59	2.24	0.136
	Eng x Site	9.10	4	2.28	1.11	0.352
	Error	454.61	222	2.05		
S-intrcp	English	359.71	1	359.71	52.58	0.001
	Spanish	2.21	1	2.21	0.32	0.570
	Site	49.76	4	12.44	1.82	0.126
	Eng x Spn	4.84	1	4.84	0.71	0.401
	Eng x Site	28.21	4	7.05	1.03	0.392
	Error	1518.74	222	6.84		
I-intrcp	English	213.80	1	213.80	46.05	0.001
	Spanish	17.00	1	17.00	3.66	0.057
	Site	62.92	4	15.73	3.39	0.010
	Eng x Spn	7.85	1	7.85	1.69	0.195
	Eng x Site	22.06	4	5.52	1.19	0.317
	Error	1030.79	222	4.64		

Table [illegible]

Interactive Reading Assessment System - English:

ANOVA Summary Table for Expository Listening Comprehension Growth Indices for the Bilingual Sample

Growth Index	Variable	Sum of Squares	df	Mean Square	F	P
Slope	English	44.47	1	44.47	16.32	0.001
	Spanish	10.61	1	10.61	3.89	0.050
	Site	47.64	4	11.91	4.37	0.002
	Eng x Spn	9.82	1	9.82	3.61	0.059
	Eng x Site	4.72	4	1.18	0.43	0.785
	Error	599.38	220	2.72		
S-intrcp	English	640.36	1	640.36	65.34	0.001
	Spanish	13.83	1	13.83	1.41	0.236
	Site	136.35	4	34.09	3.48	0.009
	Eng x Spn	1.59	1	1.59	0.16	0.688
	Eng x Site	42.39	4	10.60	1.08	0.367
	Error	2156.19	220	9.80		
I-intrcp	English	183.92	1	183.92	42.12	0.001
	Spanish	0.62	1	0.62	0.14	0.707
	Site	14.98	4	3.75	0.86	0.490
	Eng x Spn	0.78	1	0.78	0.18	0.672
	Eng x Site	36.93	4	9.23	2.11	0.080
	Error	960.67	220	4.37		

APPENDIX D

Interactive Reading Assessmen System - Spanish:

Descriptive Statistics on G owth Indices
for the Bilingual Sample for Each Site
Overall and by Language Category

Table 1

Interactive Reading Assessment System - Spanish:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Non-comprehension Scales for Site 0 Overall by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-0	Lo Hi-0	Hi Lo-0	Hi Hi-0
VDCS	Slope	M	2.5	1.9	2.5		2.7
VDCS	Slope	S	2.0	2.2	1.9		2.1
VDCS	Slope	N	51	9	13		29
VDCS	S-Intrcp	M	1.2	-1.4	0.5		2.4
VDCS	S-Intrcp	S	4.8	1.7	3.0		5.6
VDCS	I-Intrcp	M	0.6	1.4	1.0		0.2
VDCS	I-Intrcp	S	1.8	1.0	1.1		2.1
VDFS	Slope	M	1.7	2.2	1.[illegible]		1.4
VDFS	Slope	S	1.6	2.2	1.2		1.2
VDFS	Slope	N	51	9	13		29
VDFS	S-Intrcp	M	5.3	2.7	4.8		6.3
VDFS	S-Intrcp	S	4.1	4.4	3.4		4.0
VDFS	I-Intrcp	M	-1.8	-0.1	-2.0		-2.2
VDFS	I-Intrcp	S	2.3	1.5	2.4		2.4
LDCS	Slope	M	1.1	0.9	1.7		0.9
LDCS	Slope	S	1.1	0.9	1.4		1.0
LDCS	Slope	N	51	9	13		29
LDCS	S-Intrcp	M	0.1	-1.0	-1.4		1.2
LDCS	S-Intrcp	S	3.2	1.6	3.2		3.3
LDCS	I-Intrcp	M	.0	1.2	1.4		-1.1
LDCS	I-Intrcp	S	2.9	1.4	2.2		3.2
LSPS	Slope	M	11.4	7.9	15.5		10.6
LSPS	Slope	S	10.5	8.0	14.6		8.6
LSPS	Slope	N	51	9	13		29
LSPS	S-Intrcp	M	6.7	-4.3	-4.2		15.0
LSPS	S-Intrcp	S	31.0	17.6	23.2		35.1
LSPS	I-Intrcp	M	0.2	0.9	0.7		-0.3
LSPS	I-Intrcp	S	2.2	1.7	2.0		2.3
SRDS	Slope	M	0.6	0.4	0.7		0.6
SRDS	Slope	S	0.8	0.5	0.9		0.7
SRDS	Slope	N	49	8	13		28
SRDS	S-Intrcp	M	0.3	-0.6	-0.2		0.8
SRDS	S-Intrcp	S	2.4	0.7	2.7		2.4
SRDS	I-Intrcp	M	0.8	1.6	1.1		0.4
SRDS	I-Intrcp	S	1.9	1.5	1.5		2.2

511

Table 2

Interactive Reading Assessment System - Spanish:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Comprehension Scales for Site 0 Overall and by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-0	Lo Hi-0	Hi Lo-0	Hi Hi-0
NRCS	Slope	M	1.2	0.0	1.6		1.4
NRCS	Slope	S	1.3	0.0	1.9		1.1
NRCS	Slope	N	51	9	13		29
NRCS	S-Intrcp	M	-1.0	0.0	-2.4		-0.7
NRCS	S-Intrcp	S	3.1	0.0	4.5		2.6
NRCS	I-Intrcp	M	0.8	0.0	1.2		0.8
NRCS	I-Intrcp	S	1.4	0.0	1.3		1.5
ERCS	Slope	M	1.6	0.0	1.6		2.1
ERCS	Slope	S	1.8	0.0	2.1		1.8
ERCS	Slope	N	49	8	13		28
ERCS	S-Intrcp	M	-2.4	0.0	-3.0		-2.8
ERCS	S-Intrcp	S	4.5	0.0	5.0		4.8
ERCS	I-Intrcp	M	1.0	0.0	1.2		1.2
ERCS	I-Intrcp	S	2.0	0.0	1.4		2.4
NLCS	Slope	M	1.1	0.9	1.1		1.1
NLCS	Slope	S	0.8	1.1	0.7		0.8
NLCS	Slope	N	51	9	13		29
NLCS	S-Intrcp	M	2.8	2.3	2.5		3.1
NLCS	S-Intrcp	S	2.0	2.4	1.9		1.9
NLCS	I-Intrcp	M	-1.7	-0.7	-1.9		-2.0
NLCS	I-Intrcp	S	2.2	1.9	2.3		2.3
ELCS	Slope	M	1.4	1.2	1.6		1.4
ELCS	Slope	S	1.3	1.7	0.9		1.3
ELCS	Slope	N	49	8	13		28
ELCS	S-Intrcp	M	1.3	.0	0.7		2.1
ELCS	S-Intrcp	S	4.0	4.8	2.6		4.2
ELCS	I-Intrcp	M	-0.7	-0.2	0.3		-1.4
ELCS	I-Intrcp	S	2.8	3.3	1.8		2.9

Table 3

Interactive Reading Assessment System - Spanish:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Non-comprehension Scales for Site 1 Overall by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-1	Lo Hi-1	Hi Lo-1	Hi Hi-1
VDCS	Slope	M	2.7	1.6	3.9	2.1	2.6
VDCS	Slope	S	2.2	1.3	2.9	2.0	1.5
VDCS	Slope	N	17	4	6	4	3
VDCS	S-Intrcp	M	1.7	-1.0	4.6	1.0	0.5
VDCS	S-Intrcp	S	5.4	1.3	7.8	5.1	1.4
VDCS	I-Intrcp	M	0.6	1.5	-0.2	1.1	0.5
VDCS	I-Intrcp	S	1.6	0.4	2.4	1.0	1.6
VDFS	Slope	M	1.8	1.4	1.4	2.4	2.4
VDFS	Slope	S	1.3	0.7	1.1	1.8	1.8
VDFS	Slope	N	17	[illegible]	6	4	3
VDFS	S-Intrcp	M	4.2	[illegible]	7.4	1.5	3.2
VDFS	S-Intrcp	S	3.6	1.[illegible]	3.4	2.8	3.2
VDFS	I-Intrcp	M	-1.6	-1.5	-2.5	-0.7	-1.3
VDFS	I-Intrcp	S	2.3	2.4	2.0	2.5	3.2
LDCS	Slope	M	0.9	0.8	0.5	1.3	1.3
LDCS	Slope	S	0.8	0.7	0.6	1.3	0.5
LDCS	Slope	N	17	4	6	4	3
LDCS	S-Intrcp	M	1.3	-0.4	3.1	0.9	0.7
LDCS	S-Intrcp	S	2.2	0.6	2.0	2.6	1.4
LDCS	I-Intrcp	M	-0.6	1.3	-1.6	-1.5	0.1
LDCS	I-Intrcp	S	2.6	0.9	2.7	3.7	1.3
LSPS	Slope	M	12.6	6.8	16.7	12.4	12.3
LSPS	Slope	S	10.1	5.3	8.8	14.0	12.7
LSPS	Slope	N	17	4	6	4	3
LSPS	S-Intrcp	M	10.2	-4.0	17.9	11.1	12.7
LSPS	S-Intrcp	S	22.9	11.9	25.1	16.0	37.8
LSPS	I-Intrcp	M	-0.5	0.9	-0.9	-1.0	-0.8
LSPS	I-Intrcp	S	2.5	1.7	2.5	2.6	3.7
SRDS	Slope	M	0.6	0.7	0.6	0.5	0.7
SRDS	Slope	S	0.4	0.3	0.2	0.7	0.2
SRDS	Slope	N	17	4	6	4	3
SRDS	S-Intrcp	M	-0.1	-0.7	0.5	-0.1	-0.3
SRDS	S-Intrcp	S	0.8	0.6	0.2	1.3	0.2
SRDS	I-Intrcp	M	1.0	1.8	0.1	1.3	1.5
SRDS	I-Intrcp	S	0.9	0.7	0.2	1.0	0.2

Table 4

Interactive Reading Assessment System - Spanish:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Comprehension Scales for Site 1 Overall and by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-1	Lo Hi-1	Hi Lo-1	Hi Hi-1
NRCS	Slope	M	2.2	2.1	2.7	1.5	2.1
NRCS	Slope	S	1.3	2.0	0.7	1.5	0.6
NRCS	Slope	N	17	4	6	4	3
NRCS	S-Intrcp	M	-3.5	-5.7	-2.9	-2.5	-2.8
NRCS	S-Intrcp	S	3.7	5.5	3.9	2.7	1.9
NRCS	I-Intrcp	M	2.2	2.8	1.9	2.0	2.3
NRCS	I-Intrcp	S	1.3	1.8	1.0	1.6	0.7
ERCS	Slope	M	1.9	1.4	2.9	1.6	0.9
ERCS	Slope	S	1.5	2.2	0.5	1.3	1.6
ERCS	Slope	N	17	4	6	4	3
ERCS	S-Intrcp	M	-3.3	-3.8	-4.0	-3.0	-1.6
ERCS	S-Intrcp	S	3.4	5.9	2.8	2.2	2.8
ERCS	I-Intrcp	M	1.9	1.8	2.3	2.3	0.9
ERCS	I-Intrcp	S	1.5	2.1	0.9	1.6	1.6
NLCS	Slope	M	1.1	0.8	0.7	1.6	1.4
NLCS	Slope	S	0.9	0.6	0.6	1.2	1.3
NLCS	Slope	N	17	4	6	4	3
NLCS	S-Intrcp	M	2.4	2.0	4.6	-0.3	2.2
NLCS	S-Intrcp	S	2.7	2.5	1.4	1.3	3.3
NLCS	I-Intrcp	M	-1.4	-1.9	-2.9	1.2	-1.3
NLCS	I-Intrcp	S	2.6	2.8	2.1	0.8	3.3
ELCS	Slope	M	1.4	1.3	1.2	1.6	1.6
ELCS	Slope	S	1.0	1.1	0.6	1.1	1.5
ELCS	Slope	N	17	4	6	4	3
ELCS	S-Intrcp	M	0.3	-0.3	2.3	-1.5	-0.3
ELCS	S-Intrcp	S	3.2	4.4	2.5	1.9	4.0
ELCS	I-Intrcp	M	-0.3	-1.1	-1.6	1.4	1.2
ELCS	I-Intrcp	S	2.9	4.5	2.5	1.1	1.2

Table 5

Interactive Reading Assessment System - Spanish:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Non-comprehension Scales for Site 2 Overall by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-2	Lo Hi-2	Hi Lo-2	Hi Hi-2
VDCS	Slope	M	2.4	0.7	3.1	1.6	2.3
VDCS	Slope	S	2.9	0.8	3.9	0.8	2.0
VDCS	Slope	N	36	3	16	4	13
VDCS	S-Intrcp	M	-1.8	-0.7	-1.4	-0.8	-2.9
VDCS	S-Intrcp	S	4.2	0.8	5.4	0.9	3.5
VDCS	I-Intrcp	M	1.4	1.2	0.7	1.7	2.1
VDCS	I-Intrcp	S	1.7	1.1	2.3	0.8	0.7
VDFS	Slope	M	2.0	1.5	2.5	1.6	1.6
VDFS	Slope	S	2.0	0.7	2.6	1.5	1.5
VDFS	Slope	N	35	3	15	4	13
VDFS	S-Intrcp	M	1.2	0.4	1.6	-0.3	1.5
VDFS	S-Intrcp	S	4.0	2.9	4.8	4.7	3.3
VDFS	I-Intrcp	M	-0.1	0.2	-0.2	-0.5	.0
VDFS	I-Intrcp	S	2.3	2.1	2.4	3.5	2.1
LDCS	Slope	M	1.2	0.4	1.2	0.8	1.5
LDCS	Slope	S	1.2	0.6	1.4	0.6	1.1
LDCS	Slope	N	36	3	16	4	13
LDCS	S-Intrcp	M	-1.1	-0.2	-0.8	-0.4	-1.9
LDCS	S-Intrcp	S	2.1	0.4	2.5	0.6	1.8
LDCS	I-Intrcp	M	1.4	0.5	1.0	1.7	1.8
LDCS	I-Intrcp	S	1.7	1.0	2.3	0.8	1.0
LSPS	Slope	M	9.1	2.8	12.3	8.7	6.7
LSPS	Slope	S	12.9	2.4	18.1	8.9	4.9
LSPS	Slope	N	36	3	16	4	13
LSPS	S-Intrcp	M	-1.5	-0.6	-2.4	-0.6	-1.1
LSPS	S-Intrcp	S	15.5	2.2	22.2	9.5	7.6
LSPS	I-Intrcp	M	0.5	0.8	0.7	-0.3	0.5
LSPS	I-Intrcp	S	2.0	0.9	2.1	2.5	2.1
SRDS	Slope	M	0.7	0.2	0.7	0.5	0.8
SRDS	Slope	S	0.5	0.2	0.4	0.4	0.6
SRDS	Slope	N	35	3	15	4	13
SRDS	S-Intrcp	M	-0.6	.0	-0.3	-0.5	-1.1
SRDS	S-Intrcp	S	1.2	0.8	1.0	0.4	1.4
SRDS	I-Intrcp	M	1.4	-0.8	1.2	1.4	2.0
SRDS	I-Intrcp	S	1.6	3.9	1.4	1.0	1.0

515

Table 6

Interactive Reading Assessment System - Spanish:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Comprehension Scales for Site 2 Overall and by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-2	Lo Hi-2	Hi Lo-2	Hi Hi-2
NRCS	Slope	M	1.9	1.0	1.7	1.4	2.5
NRCS	Slope	S	1.6	1.7	1.2	1.9	1.9
NRCS	Slope	N	36	3	16	4	13
NRCS	S-Intrcp	M	-3.4	-2.6	-2.3	-3.4	-4.9
NRCS	S-Intrcp	S	3.8	4.5	3.0	5.0	4.2
NRCS	I-Intrcp	M	1.8	1.2	1.6	1.6	2.3
NRCS	I-Intrcp	S	1.5	2.1	1.4	1.9	1.4
ERCS	Slope	M	1.6	0.9	1.4	1.2	2.0
ERCS	Slope	S	1.4	1.6	1.4	1.5	1.4
ERCS	Slope	N	35	3	15	4	13
ERCS	S-Intrcp	M	-2.7	-2.4	-1.6	-2.7	-4.1
ERCS	S-Intrcp	S	3.1	4.2	2.7	3.8	3.0
ERCS	I-Intrcp	M	1.7	1.2	1.3	1.5	2.4
ERCS	I-Intrcp	S	1.5	2.1	1.3	1.9	1.4
NLCS	Slope	M	1.1	0.5	1.3	1.0	1.0
NLCS	Slope	S	0.9	0.8	1.1	0.8	0.6
NLCS	Slope	N	35	3	15	4	13
NLCS	S-Intrcp	M	1.6	1.7	2.2	0.1	1.3
NLCS	S-Intrcp	S	2.0	1.9	2.3	1.0	1.8
NLCS	I-Intrcp	M	-0.8	-0.5	-1.5	0.7	-0.7
NLCS	I-Intrcp	S	2.2	2.3	2.5	0.8	2.1
ELCS	Slope	M	1.2	0.7	1.0	1.7	1.4
ELCS	Slope	S	1.4	1.0	1.9	1.2	0.8
ELCS	Slope	N	35	3	15	4	13
ELCS	S-Intrcp	M	0.1	-0.1	1.3	-2.5	-0.4
ELCS	S-Intrcp	S	3.2	3.4	3.9	2.6	2.0
ELCS	I-Intrcp	M	0.3	-2.1	-0.1	1.7	0.8
ELCS	I-Intrcp	S	2.5	4.4	2.6	1.4	2.1

Table 7

Interactive Reading Assessment System - Spanish:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Non-comprehension Scales for Site 3 Overall by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-3	Lo Hi-3	Hi Lo-3	Hi Hi-3
VDCS	Slope	M	0.7	0.2		0.5	2.7
VDCS	Slope	S	1.3	0.4		0.8	2.8
VDCS	Slope	N	74	27		38	9
VDCS	S-Intrcp	M	0.0	-0.1		-0.1	0.7
VDCS	S-Intrcp	S	1.1	0.2		0.7	2.7
VDCS	I-Intrcp	M	0.5	0.5		0.5	0.7
VDCS	I-Intrcp	S	1.0	0.8		1.2	0.9
VDFS	Slope	M	1.0	0.4		1.2	1.6
VDFS	Slope	S	3.0	2.9		2.9	4.0
VDFS	Slope	N	73	27		37	9
VDFS	S-Intrcp	M	1.2	0.7		1.1	3.1
VDFS	S-Intrcp	S	3.6	3.1		3.4	5.4
VDFS	I-Intrcp	M	0.4	0.4		0.5	-0.1
VDFS	I-Intrcp	S	1.2	1.3		0.9	2.0
LDCS	Slope	M	0.1	0.0		0.1	0.6
LDCS	Slope	S	0.5	.0		0.3	1.2
LDCS	Slope	N	75	28		38	9
LDCS	S-Intrcp	M	0.1	0.0		0.1	0.8
LDCS	S-Intrcp	S	0.9	.0		0.7	2.1
LDCS	I-Intrcp	M	0.1	0.1		0.3	-0.8
LDCS	I-Intrcp	S	1.0	0.4		0.7	2.2
LSPS	Slope	M	1.0	0.0		-0.2	10.0
LSPS	Slope	S	5.5	0.0		2.8	12.9
LSPS	Slope	N	71	28		35	8
LSPS	S-Intrcp	M	1.2	0.0		0.9	6.6
LSPS	S-Intrcp	S	6.9	0.0		4.2	18.5
LSPS	I-Intrcp	M	.0	0.0		.0	-0.3
LSPS	I-Intrcp	S	0.6	0.0		0.1	2.0
SRDS	Slope	M	0.1	.0		0.1	0.4
SRDS	Slope	[illegible]	0.3	0.1		0.3	0.5
SRDS	Slope	N	75	28		38	9
SRDS	S-Intrcp	M	.0	.0		.0	0.2
SRDS	S-Intrcp	S	0.4	0.1		0.4	0.5
SRDS	I-Intrcp	M	0.1	0.1		0.2	-0.2
SRDS	I-Intrcp	S	0.9	0.4		1.1	1.3

Table 8

Interactive Reading Assessment System - Spanish:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Comprehension Scales for Site 3 Overall and by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-3	Lo Hi-3	Hi Lo-3	Hi Hi-3
NRCS	Slope	M	0.1	0.0		0.0	0.6
NRCS	Slope	S	0.5	0.0		0.0	1.4
NRCS	Slope	N	75	28		38	9
NRCS	S-Intrcp	M	.0	0.0		0.0	-0.3
NRCS	S-Intrcp	S	0.3	0.0		0.0	0.9
NRCS	I-Intrcp	M	.0	0.0		0.0	0.3
NRCS	I-Intrcp	S	0.2	0.0		0.0	0.6
ERCS	Slope	M	0.1	0.0		0.0	0.7
ERCS	Slope	S	0.5	0.0		0.0	1.5
ERCS	Slope	N	75	28		38	9
ERCS	S-Intrcp	M	-0.1	0.0		0.0	-0.5
ERCS	S-Intrcp	S	0.4	0.0		0.0	1.0
ERCS	I-Intrcp	M	.0	0.0		0.0	0.4
ERCS	I-Intrcp	S	0.3	0.0		0.0	0.7
NLCS	Slope	M	0.2	0.2		0.3	-0.3
NLCS	Slope	S	1.1	1.0		1.2	1.1
NLCS	Slope	N	75	28		38	9
NLCS	S-Intrcp	M	0.7	0.2		0.6	2.6
NLCS	S-Intrcp	S	1.8	1.2		1.9	1.9
NLCS	I-Intrcp	M	-0.1	0.5		-0.1	-1.8
NLCS	I-Intrcp	S	1.8	1.3		1.7	2.1
ELCS	Slope	M	0.1	0.1		0.1	-0.1
ELCS	Slope	S	0.9	0.4		1.1	1.2
ELCS	Slope	N	75	28		38	9
ELCS	S-Intrcp	M	0.3	-0.1		0.2	1.8
ELCS	S-Intrcp	S	1.3	0.3		1.4	2.1
ELCS	I-Intrcp	M	0.1	0.2		0.3	-1.0
ELCS	I-Intrcp	S	1.0	0.6		0.6	1.9

Table 9

Interactive Reading Assessment System - Spanish:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Non-comprehension Scales for Site 5 Overall by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-3	Lo Hi-3	Hi Lo-3	Hi Hi-3
VDCS	Slope	M	3.3	4.3	3.5	3.5	2.5
VDCS	Slope	S	3.7	4.0	3.6	5.1	3.4
VDCS	Slope	N	69	15	19	7	28
VDCS	S-Intrcp	M	0.6	-1.3	2.4	0.8	0.3
VDCS	S-Intrcp	S	4.9	5.1	5.6	7.6	3.1
VDCS	I-Intrcp	M	0.3	0.9	0.2	0.5	0.0
VDCS	I-Intrcp	S	1.8	1.7	1.7	1.5	2.1
VDFS	Slope	M	2.6	3.8	1.8	3.4	2.2
VDFS	Slope	S	3.1	2.8	2.3	3.9	3.3
VDFS	Slope	N	69	15	19	7	28
VDFS	S-Intrcp	M	2.7	0.3	4.9	1.4	2.9
VDFS	S-Intrcp	S	4.9	4.8	3.9	6.9	4.7
VDFS	I-Intrcp	M	-0.6	0.6	-1.6	-0.7	-0.5
VDFS	I-Intrcp	S	2.4	1.8	2.5	3.0	2.2
LDCS	Slope	M	1.5	2.6	1.3	2.1	1.0
LDCS	Slope	S	1.8	1.9	1.6	1.7	1.7
LDCS	Slope	N	64	14	19	7	24
LDCS	S-Intrcp	M	0.4	-1.3	1.5	-0.6	0.8
LDCS	S-Intrcp	S	2.9	2.1	3.1	2.5	2.8
LDCS	I-intrcp	M	0.2	1.2	-0.7	0.6	0.2
LDCS	I-Intrcp	S	2.1	1.1	2.5	2.5	1.9
LSPS	Slope	M	16.3	21.5	19.2	18.1	11.2
LSPS	Slope	S	19.3	19.5	13.9	20.3	21.7
LSPS	Slope	N	69	15	19	7	28
LSPS	S-Intrcp	M	6.2	-6.1	12.8	2.9	9.2
LSPS	S-Intrcp	S	29.9	29.2	28.6	31.4	30.2
LSPS	I-Intrcp	M	0.2	0.9	-0.5	-0.3	0.5
LSPS	I-Intrcp	S	1.8	1.8	2.2	2.3	1.3
SRDS	Slope	M	0.7	0.7	1.0	0.8	0.6
SRDS	Slope	S	0.6	0.4	0.8	0.5	0.5
SRDS	Slope	N	69	15	19	7	28
SRDS	S-Intrcp	M	-0.1	-0.3	-0.1	-0.2	-0.1
SRDS	S-Intrcp	S	0.5	0.5	0.5	0.5	0.4
SRDS	I-Intrcp	M	0.7	1.0	0.6	1.2	0.6
SRDS	I-Intrcp	S	1.4	1.7	1.4	0.7	1.4

519

Table . 10

Interactive Reading Assessment System - Spanish:
Descriptive Statistics on Growth Indices for the Bilingual Sample
Comprehension Scales for Site 5 Overall and by Language Category

Scale	Measure	Statistic	Overall	Lo Lo-3	Lo Hi-3	Hi Lo-3	Hi Hi-3
NRCS	Slope	M	0.8	0.4	1.3	1.0	0.5
NRCS	Slope	S	1.3	0.9	1.6	1.3	1.2
NRCS	Slope	N	69	15	19	[illegible]	28
NRCS	S-Intrcp	M	-0.3	-0.3	-0.3	-0.5	-0.3
NRCS	S-Intrcp	S	1.1	0.7	1.7	0.9	0.9
NRCS	I-Intrcp	M	0.4	0.3	0.6	0.6	0.3
NRCS	I-Intrcp	S	0.7	0.7	0.9	0.8	0.6
ERCS	Slope	M	0.7	0.2	1.2	1.2	0.5
ERCS	Slope	S	1.3	0.7	1.7	1.5	1.2
ERCS	Slope	N	69	15	19	7	28
ERCS	S-Intrcp	M	-0.4	-0.1	-0.3	-0.8	-0.4
ERCS	S-Intrcp	S	1.2	0.5	1.8	1.1	0.8
ERCS	I-Intrcp	M	0.3	0.2	0.3	0.7	0.3
ERCS	I-Intrcp	S	1.0	0.6	1.6	0.9	0.7
NLCS	Slope	M	1.2	1.3	1.2	0.5	1.4
NLCS	Slope	S	1.3	1.4	1.6	1.5	1.0
NLCS	Slope	N	68	15	18	7	28
NLCS	S-Intrcp	M	1.9	1.4	2.3	1.9	1.9
NLCS	S-Intrcp	S	2.5	2.6	2.5	2.6	2.6
NLCS	I-Intrcp	M	-0.8	-0.5	-1.2	-0.4	-0.7
NLCS	I-Intrcp	S	2.3	2.6	2.4	2.1	2.3
ELCS	Slope	M	1.1	0.7	1.4	0.6	1.1
ELCS	Slope	S	1.8	2.0	1.8	2.5	1.6
ELCS	Slope	N	69	15	19	7	28
ELCS	S-Intrcp	M	1.2	1.1	1.3	0.4	1.2
ELCS	S-Intrcp	S	2.5	2.8	2.5	2.7	2.3
ELCS	I-Intrcp	M	-0.1	0.7	-0.7	1.3	-0.4
ELCS	I-Intrcp	S	2.0	0.8	2.2	0.6	2.2

530

APPENDIX E

Interactive Reading Assessment System - Spanish:
Summary Tables for the Analyses of Variance on Growth Indices for the Bilingual Sample

Table 1

Interactive Reading Assessment System - Spanish:

ANOVA Summary Table for Real Word Decoding Growth Indices
for the Bilingual Sample

Growth Index	Variable	Sum of Squares	df	Mean Square	F	P
Slope	English	4.19	1	4.19	0.61	0.434
	Spanish	15.73	1	15.73	2.31	0.130
	Site	138.03	4	34.51	5.06	0.001
	Eng x Spn	4.01	1	4.01	0.59	0.444
	Spn x Site	48.01	4	12.00	1.76	0.138
	Error	1528.76	224	6.83		
S-intrcp	English	0.60	1	0.60	0.04	0.85
	Spanish	97.16	1	97.16	6.10	0.014
	Site	268.55	4	67.14	4.21	0.003
	Eng x Spn	78.64	1	78.64	4.94	0.027
	Spn x Site	65.38	4	16.34	1.03	0.395
	Error	3569.54	224	15.94		
I-intrcp	English	.00	1	.00	.00	0.980
	Spanish	11.55	1	11.55	4.61	0.033
	Site	35.56	4	8.89	3.55	0.008
	Eng x Spn	1.76	1	1.76	0.70	0.402
	Spn x Site	3.09	4	0.77	0.31	0.872
	Error	561.32	224	2.51		

522

Table 2

Interactive Reading Assessment System - Spanish:

ANOVA Summary Table for Vocabulary Definition Growth Indices for the Bilingual Sample

Growth Index	Variable	Sum of Squares	df	Mean Square	F	p
Slope	English	0.48	1	0.48	0.07	0.789
	Spanish	12.96	1	12.96	1.97	0.162
	Site	100.76	4	25.19	3.83	0.005
	Eng x Spn	1.24	1	1.24	0.19	0.664
	Spn x Site	30.59	4	7.65	1.16	0.328
	Error	1461.01	222	6.58		
S-intrcp	English	0.42	1	0.42	0.03	0.873
	Spanish	298.30	1	298.30	18.21	0.001
	Site	411.69	4	102.92	6.28	0.001
	Eng x Spn	47.89	1	47.89	2.92	0.089
	Spn x Site	41.29	4	10.32	0.63	0.641
	Error	3635.89	222	16.38		
I-intrcp	English	0.35	1	0.35	0.08	0.773
	Spanish	39.99	1	39.99	9.63	0.002
	Site	95.26	4	23.82	5.73	0.001
	Eng x Spn	22.70	1	22.70	5.46	0.020
	Spn x Site	9.86	4	2.47	0.59	0.668
	Error	922.30	222	4.15		

Table 3

Interactive Reading Assessment System - Spanish:

ANOVA Summary Table for Letter-Sound Decoding Growth Indices
for the Bilingual Sample

Growth Index	Variable	Sum of Squares	df	Mean Square	F	p
Slope	English	0.65	1	0.65	0.47	0.495
	Spanish	1.81	1	1.81	1.29	0.257
	Site	63.19	4	15.80	11.31	0.001
	Eng x Spn	0.33	1	0.33	0.24	0.626
	Spn x Site	29.95	4	7.49	5.36	0.001
	Error	307.38	220	1.40		
S-intrcp	English	2.44	1	2.44	0.48	0.49
	Spanish	54.41	1	54.41	10.79	0.001
	Site	116.94	4	29.23	5.80	0.001
	Eng x Spn	17.42	1	17.42	3.45	0.064
	Spn x Site	70.92	4	17.73	3.52	0.008
	Error	1109.64	220	5.04		
I-intrcp	English	2.30	1	2.30	0.63	0.429
	Spanish	32.79	1	32.79	8.92	0.003
	Site	82.33	4	20.58	5.60	0.001
	Eng x Spn	19.08	1	19.08	5.19	0.024
	Spn x Site	37.31	4	9.33	2.54	0.041
	Error	808.26	220	3.67		

Table 4

Interactive Reading Assessment System - Spanish:

ANOVA Summary Table for Letter-Sound Spelling Growth Indices
for the Bilingual Sample

Growth Index	Variable	Sum of Squares	df	Mean Square	F	P
Slope	English	504.73	1	504.73	3.02	0.084
	Spanish	201.66	1	201.66	1.21	0.273
	Site	5342.98	4	1335.74	8.00	0.001
	Eng x Spn	262.07	1	262.07	1.57	0.212
	Spn x Site	1196.22	4	299.05	1.79	0.132
	Error	36918.96	221	167.05		
S-intrcp	English	906.26	1	906.26	1.71	0.193
	Spanish	2943.71	1	2943.71	5.55	0.019
	Site	2753.73	4	688.43	1.30	0.272
	Eng x Spn	667.80	1	667.80	1.26	0.263
	Spn x Site	1739.44	4	434.86	0.82	0.514
	Error	117214.18	221	530.38		
I-intrcp	English	2.31	1	2.31	0.75	0.387
	Spanish	4.88	1	4.88	1.59	0.208
	Site	18.33	4	4.58	1.50	0.205
	Eng x Spn	20.03	1	20.03	6.53	0.011
	Spn x Site	11.76	4	2.94	0.96	0.431
	Error	677.54	221	3.07		

Table 5

Interactive Reading Assessment System - Spanish:

ANOVA Summary Table for Sentence Reading Growth Indices
for the Bilingual Sample

Growth Index	Variable	Sum of Squares	df	Mean Square	F	p
Slope	English	0.18	1	0.18	0.64	0.426
	Spanish	1.09	1	1.09	3.85	0.051
	Site	7.81	4	1.95	6.90	0.001
	Eng x Spn	0.71	1	0.71	2.49	0.116
	Spn x Site	0.76	4	0.19	0.67	0.614
	Error	62.87	222	0.28		
S-intrcp	English	0.66	1	0.66	0.47	0.496
	Spanish	3.05	1	3.05	2.14	0.145
	Site	19.91	4	4.98	3.49	0.009
	Eng x Spn	0.52	1	0.52	0.36	0.547
	Spn x Site	1.97	4	0.49	0.35	0.847
	Error	316.53	222	1.43		
I-intrcp	English	0.23	1	0.23	0.12	0.730
	Spanish	5.19	1	5.19	2.65	0.105
	Site	45.80	4	11.45	5.85	0.001
	Eng x Spn	0.16	1	0.16	0.08	0.773
	Spn x Site	13.29	4	3.32	1.70	0.152
	Error	434.53	222	1.96		

528

Table 5

Interactive Reading Assessment System - Spanish:

ANOVA Summary Table for Narrative Reading Comprehension Growth Indices for the Bilingual Sample

Growth Index	Variable	Sum of Squares	df	Mean Square	F	P
Slope	English	0.13	1	0.13	0.11	0.747
	Spanish	16.60	1	16.60	13.03	0.001
	Site	67.42	4	16.85	13.23	0.001
	Eng x Spn	4.22	1	4.22	3.31	0.070
	Spn x Site	6.94	4	1.74	1.36	0.248
	Error	286.66	225	1.27		
S-intrcp	English	0.09	1	0.09	0.02	0.90
	Spanish	1.38	1	1.38	0.27	0.602
	Site	356.98	4	89.25	17.65	0.001
	Eng x Spn	2.28	1	2.28	0.45	0.503
	Spn x Site	48.12	4	12.03	2.38	0.053
	Error	1137.52	225	5.06		
I-intrcp	English	0.02	1	0.02	0.02	0.888
	Spanish	2.46	1	2.46	2.70	0.102
	Site	97.77	4	24.44	26.87	0.001
	Eng x Spn	0.39	1	0.39	0.43	0.512
	Spn x Site	5.79	4	1.45	1.59	0.178
	Error	204.66	225	0.91		

Table 7

Interactive Reading Assessment System - Spanish:

ANOVA Summary Table for Expository Reading Comprehension Growth Indices for the Bilingual Sample

Growth Index	Variable	Sum of Squares	df	Mean Square	F	p
Slope	English	0.43	1	0.43	0.29	0.590
	Spanish	19.90	1	19.90	13.48	0.001
	Site	51.59	4	12.90	8.74	0.001
	Eng x Spn	12.90	1	12.90	8.74	0.003
	Spn x Site	8.74	4	2.18	1.48	0.209
	Error	327.72	222	1.48		
S-intrcp	English	7.34	1	7.34	1.18	0.279
	Spanish	9.00	1	9.00	1.44	0.231
	Site	283.27	4	70.72	11.36	0.001
	Eng x Spn	2.16	1	2.16	0.35	0.557
	Spn x Site	28.83	4	7.21	1.16	0.331
	Error	1383.75	222	6.23		
I-intrcp	English	2.26	1	2.26	1.55	0.214
	Spanish	1.72	1	1.72	1.18	0.278
	Site	85.52	4	21.38	14.68	0.001
	Eng x Spn	2.96	1	2.96	2.03	0.155
	Spn x Site	6.53	4	1.63	1.12	0.348
	Error	323.28	222	1.46		

Table 9

Interactive Reading Assessment System - Spanish:

ANOVA Summary Table for Narrative Listening Comprehension Growth Indices for the Bilingual Sample

Growth Index	Variable	Sum of Squares	df	Mean Square	F	p
Slope	English	.00	1	.00	.00	0.965
	Spanish	0.22	1	0.22	0.19	0.664
	Site	29.17	4	7.29	6.21	0.001
	Eng x Spn	0.55	1	0.55	0.47	0.493
	Spn x Site	5.84	4	1.46	1.24	0.294
	Error	261.97	223	1.18		
S-intrcp	English	2.47	1	2.47	0.59	0.444
	Spanish	58.49	1	58.49	13.92	0.001
	Site	51.95	4	12.99	3.09	0.017
	Eng x Spn	4.46	1	4.46	1.06	0.304
	Spn x Site	37.87	4	9.47	2.25	0.064
	Error	937.11	223	4.20		
I-intrcp	English	2.21	1	2.21	0.53	0.468
	Spanish	57.53	1	57.53	13.75	0.001
	Site	32.26	4	8.07	1.93	0.107
	Eng x Spn	1.80	1	1.80	0.43	0.512
	Spn x Site	12.88	4	3.22	0.77	0.546
	Error	933.09	223	4.18		

Table 9

Interactive Reading Assessment System - Spanish:

ANOVA Summary Table for Expository Listening Comprehension Growth Indices
for the Bilingual Sample

Growth Index	Variable	Sum of Squares	df	Mean Square	F	p
Slope	English	0.03	1	0.03	0.02	0.893
	Spanish	0.76	1	0.76	0.40	0.528
	Site	39.02	4	9.76	5.13	0.001
	Eng x Spn	0.95	1	0.95	0.50	0.480
	Spn x Site	1.68	4	0.42	0.22	0.927
	Error	422.33	222	1.90		
S-intrcp	English	2.21	1	2.21	0.31	0.576
	Spanish	66.89	1	66.89	9.48	0.002
	Site	40.4	4	10.11	1.43	0.224
	Eng x Spn	0.07	1	0.07	0.01	0.919
	Spn x Site	10.47	4	2.62	0.37	0.829
	Error	1566.52	222	7.06		
I-intrcp	English	1.73	1	1.73	0.45	0.503
	Spanish	37.30	1	37.30	9.72	0.002
	Site	27.17	4	6.79	1.77	0.136
	Eng x Spn	0.79	1	0.79	0.21	0.651
	Spn x Site	24.35	4	6.09	1.59	0.179
	Error	852.06	222	3.84		

530

APPENDIX F

Correlations for IRAS-English and Spanish Growth Indices

Table 1

Interactive Reading Assessment System - English and Spanish:
Correlations Between Scales for Each Growth Index
for the Bilingual Sample

	VDC SLOPE	VDF SLOPE	LDC SLOPE	LSP SLOPE	SRD SLOPE	NRC SLOPE	ERC SLOPE	NLC SLOPE	ELC SLOPE
VDC SLOPE	-	0.38	0.46	0.50	0.52	0.48	0.39	0.14	0.15
VDF SLOPE	0.40	-	0.18	0.21	0.15	0.09	0.08	0.28	0.14
LDC SLOPE	0.53	0.32	-	0.35	0.36	0.26	0.20	0.24	0.20
LSP SLOPE	0.61	0.27	.71	-	0.33	0.25	0.18	-0.02	[illegible]
SRD SLOPE	0.51	0.20	0.35	0.40	-	0.65	0.58	0.23	0.30
NRC SLOPE	0.38	0.06	0.16	[illegible]	0.56	-	0.81	0.25	0.42
ERC SLOPE	0.39	0.04	0.08	[illegible]	0.50	0.85	-	[illegible]	0.45
NLC SLOPE	0.23	0.13	0.07	0.18	0.23	0.29	0.29	-	0.61
ELC SLOPE	0.19	0.07	0.11	0.08	0.23	0.44	0.42	0.54	-

	VDC S-INT	VDF S-INT	LDC S-INT	LSP S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
VDC S-INT	-	0.41	0.58	0.56	0.67	0.57	0.45	0.31	0.41
VDF S-INT	0.49	-	0.31	0.32	0.74	0.20	0.19	0.47	0.38
LDC S-INT	0.67	0.42	-	0.56	0.45	0.27	0.25	0.34	0.33
LSP S-INT	0.66	0.42	0.76	-	0.41	0.32	0.26	0.22	0.26
SRD S-INT	0.44	0.20	0.44	0.36	-	0.59	0.48	0.28	0.36
NRC S-INT	0.22	-0.01	0.17	0.08	0.42	-	0.77	0.27	0.42
ERC S-INT	0.10	-0.13	-0.01	-0.08	0.24	0.71	-	0.36	0.46
NLC S-INT	0.15	0.40	0.09	0.17	0.06	0.05	-0.02	-	.69
ELC S-INT	0.14	0.27	0.12	0.10	0.10	0.28	0.26	0.56	-

	VDC I-INT	VDF I-INT	LDC I-INT	LSP I-INT	SRD I-INT	NRC I-INT	ERC I-INT	NLC I-INT	ELC I-INT
VDC I-INT	-	0.31	0.37	0.30	0.48	0.49	0.33	0.24	[illegible]
VDF I-INT	0.20	-	0.05	0.12	0.24	0.09	0.10	0.26	[illegible]
LDC I-INT	0.50	0.18	-	0.36	0.43	.28	0.25	0.25	.25
LSP I-INT	0.36	0.16	0.58	-	0.35	0.25	0.24	[illegible]	[illegible]
SRD I-INT	0.33	0.04	0.37	0.30	-	0.40	.39	[illegible]	[illegible]
NRC I-INT	0.16	-0.17	0.11	00	0.34	-	0.66	0.15	[illegible]
ERC I-INT	0.15	-0.19	0.03	0.02	0.24	[illegible]	-	0.09	0.51
NLC I-INT	0.07	0.36	0.04	0.08	0.02	-0.05	[illegible]	-	.48
ELC I-INT	[illegible]	0.17	0.19	00	-0.03	0.13	[illegible]	[illegible]	-

Note: English values are above the diagonal, Spanish below.

532

Table 2

Interactive Reading Assessment System - English:
Correlations Between Scales and Growth Indices
for the Bilingual Sample

	VDC S-INT	VDF S-INT	LDC S-INT	LSP S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
VDC SLOPE	-0.49	-0.07	-0.06	-0.12	-0.18	-0.17	-0.12	0.01	0.02
VDF SLOPE	-0.26	-0.81	-0.11	-0.16	-0.13	-0.05	-0.05	-0.26	-0.17
LDC SLOPE	-0.39	-0.11	-0.72	-0.33	-0.26	-0.15	-0.13	-0.18	-0.15
LSP SLOPE	-0.13	0.03	-0.04	-0.52	-0.02	.00	.00	0.13	0.11
SRD SLOPE	-0.27	-0.02	-0.10	-0.13	-0.59	-0.28	-0.21	-0.02	-0.05
NRC SLOPE	-0.23	0.07	0.03	-0.05	-0.35	-0.66	-0.48	-0.03	-0.14
ERC SLOPE	-0.12	0.06	0.04	0.02	-0.25	-0.50	-0.68	-0.04	-0.19
NLC SLOPE	-0.14	-0.26	-0.17	-0.04	-0.18	-0.17	-0.16	-0.80	-0.49
ELC SLOPE	-0.16	-0.15	-0.09	-0.02	-0.22	-0.30	-0.32	-0.47	-0.76

	VDC I-INT	VDF I-INT	LDC I-INT	LSP I-INT	SRD I-INT	NRC I-INT	ERC I-INT	NLC I-INT	ELC I-INT
VDC SLOPE	0.36	0.08	-0.03	0.04	0.17	0.24	0.15	-0.07	-0.01
VDF SLOPE	0.24	0.52	0.02	0.03	0.11	0.11	0.05	0.19	0.09
LDC SLOPE	0.29	0.06	0.47	0.23	0.24	0.16	0.14	0.15	0.13
LSP SLOPE	.00	-0.01	-0.02	0.30	0.03	0.01	0.03	-0.17	-0.13
SRD SLOPE	0.17	-0.08	0.12	0.20	0.42	0.32	0.27	-0.08	-0.01
NRC SLOPE	0.15	-0.13	0.03	0.06	0.23	0.59	0.44	-0.09	0.09
ERC SLOPE	0.06	-0.09	0.05	0.04	0.17	0.46	0.59	-0.06	0.16
NLC SLOPE	0.09	0.20	0.11	0.04	0.12	0.17	0.17	0.57	0.33
ELC SLOPE	0.06	0.08	0.09	0.09	0.14	0.28	0.36	0.33	0.54

	VDC S-INT	VDF S-INT	LDC S-INT	LSP S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
VDC I-INT	-0.81	-0.38	-0.49	-0.41	-0.53	-0.47	-0.35	-0.28	-0.30
VDF I-INT	-0.30	-0.65	-0.19	-0.23	-0.17	-0.06	-0.09	-0.32	-0.22
LDC I-INT	-0.48	-0.21	-0.77	-0.45	-0.46	-0.28	-0.30	-0.23	-0.23
LSP I-INT	-0.40	-0.22	-0.42	-0.73	-0.43	-0.29	-0.24	-0.19	-0.28
SRD I-INT	-0.54	-0.29	-0.39	-0.34	-0.74	-0.45	-0.36	-0.20	-0.25
NRC I-INT	-0.53	-0.23	-0.24	-0.29	-0.53	-0.80	-0.64	[illegible]	-0.37
ERC I-INT	-0.42	-0.18	-0.22	-0.26	-0.45	-0.61	-0.82	-0.22	-0.47
NLC I-INT	-0.28	-0.40	-0.35	-0.21	-0.21	-0.15	-0.16	-0.82	-0.58
ELC I-INT	-0.41	-0.33	-0.29	-0.21	-0.24	-0.40	-0.48	-0.53	-0.79

Table 3

Interactive Reading Assessment System - Spanish:
Correlations Between Scales and Growth Indices
for the Bilingual Sample

	VDC S-INT	VDF S-INT	LDC S-INT	LSP S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
VDC SLOPE	-0.45	-0.05	[illegible]	-0.10	-0.16	-0.16	-0.23	[illegible]	[illegible]
VDF SLOPE	-0.26	-0.68	-0.19	-0.15	-0.10	-0.03	-0.01	-0.07	-0.06
LDC SLOPE	-0.43	-0.15	-0.67	-0.48	-0.28	-0.19	-0.11	0.15	0.04
LSP SLOPE	-0.26	-0.02	-0.26	-0.51	-0.17	-0.07	-0.36	[illegible]	0.12
SRD SLOPE	[illegible]	0.15	0.02	0.09	-0.58	-0.37	-0.33	0.16	0.08
NRC SLOPE	0.09	[illegible]	0.17	0.24	-0.15	-0.78	-0.67	0.12	[illegible]
ERC SLOPE	0.16	0.30	0.28	0.33	-0.03	-0.55	-0.83	0.15	-0.09
NLC SLOPE	0.03	0.05	0.10	0.03	0.01	-0.18	-0.19	-0.51	-0.30
ELC SLOPE	0.04	[illegible]	0.07	0.09	0.01	-0.31	-0.35	-0.39	-0.64

	VDC I-INT	VDF I-INT	LDC I-INT	LSP I-INT	SRD I-INT	NRC I-INT	ERC I-INT	NLC I-INT	ELC I-INT
VDC SLOPE	0.29	-0.04	-0.01	0.11	0.22	0.25	0.24	-0.09	-0.06
VDF SLOPE	0.17	0.39	0.08	0.12	0.16	0.04	0.01	0.09	0.14
LDC SLOPE	0.26	-0.01	0.41	0.34	0.33	0.20	0.16	-0.05	0.06
LSP SLOPE	0.21	0.03	0.14	0.35	0.27	0.12	0.10	-0.04	-0.07
SRD SLOPE	0.14	-0.14	0.02	0.07	0.48	0.44	0.42	[illegible]	[illegible]
NRC SLOPE	0.05	-0.28	-0.09	-0.16	0.21	0.82	0.65	[illegible]	0.06
ERC SLOPE	-0.02	-0.30	-0.22	-0.20	0.12	0.62	0.77	-0.11	0.00
NLC SLOPE	.00	-0.02	-0.07	0.02	0.08	0.23	[illegible]	[illegible]	0.21
ELC SLOPE	0.03	-0.17	-0.04	0.02	0.01	0.37	0.34	[illegible]	0.27

	VDC S-INT	VDF S-INT	LDC S-INT	LSP S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
VDC I-INT	-0.63	-0.26	-0.43	-0.49	-0.43	-0.22	[illegible]	[illegible]	[illegible]
VDF I-INT	-0.30	-0.67	-0.25	-0.30	-0.12	0.15	0.20	[illegible]	[illegible]
LDC I-INT	-0.35	-0.30	-0.76	-0.63	-0.35	-0.20	.00	[illegible]	[illegible]
LSP I-INT	-0.47	-0.25	-0.54	-0.72	-0.36	-0.04	0.04	-0.07	-0.09
SRD I-INT	-0.33	-0.11	-0.39	-0.33	-0.55	-0.31	-0.20	[illegible]	[illegible]
NRC I-INT	-0.12	0.10	-0.07	0.02	-0.29	-0.85	-0.64	[illegible]	[illegible]
ERC I-INT	-0.06	0.14	-0.03	0.06	-0.30	-0.64	-0.36	[illegible]	[illegible]
NLC I-INT	[illegible]	-0.34	-0.08	-0.16	-0.06	-0.01	[illegible]	-0.72	[illegible]
ELC I-INT	[illegible]	-0.28	-0.17	-0.05	-0.12	-0.22	-0.14	[illegible]	[illegible]

531

Table 4

Interactive Reading Assessment System:
Correlations Between English and Spanish
Growth Indices for the Bilingual Sample
English slopes vs Spanish growth indices)

		Spanish								
		VDC SLOPE	VDF SLOPE	LDC SLOPE	LSP SLOPE	SRD SLOPE	NRC SLOPE	ERC SLOPE	NLC SLOPE	ELC SLOPE
	VDC SLOPE	0.10	0.01	-0.11	0.09	0.12	0.18	0.19	-0.03	0.03
	VDF SLOPE	0.15	0.14	0.13	0.23	0.11	0.06	0.10	0.07	0.09
	LDC SLOPE	0.16	-0.02	-0.05	0.03	0.05	0.05	0.04	-0.03	-0.02
	LSP SLOPE	0.02	-0.02	-0.11	-0.02	.00	0.02	0.03	-0.11	-0.10
English	SRD SLOPE	-0.08	-0.09	-0.21	-0.14	0.08	-0.04	-0.02	-0.09	-0.07
	NRC SLOPE	-0.04	-0.16	-0.15	-0.12	0.13	0.13	0.12	-0.06	-0.02
	ERC SLOPE	-0.05	-0.10	-0.13	-0.13	0.23	0.16	0.15	-0.12	0.04
	NLC SLOPE	-0.06	-0.15	-0.13	-0.05	0.10	-0.02	0.00	0.13	0.02
	ELC SLOPE	-0.11	-0.14	-0.14	-0.07	0.10	0.01	-0.03	-0.01	-0.02

		Spanish								
		VDC S-INT	VDF S-INT	LDC S-INT	LSP S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
	VDC SLOPE	0.18	0.07	0.04	0.21	0.10	0.01	-0.03	-0.03	0.03
	VDF SLOPE	0.11	-0.02	0.08	0.02	0.06	0.06	-0.01	0.01	0.02
	LDC SLOPE	-0.12	-0.10	-0.01	-0.04	-0.08	-0.03	-0.03	-0.17	-0.13
	LSP SLOPE	0.07	.00	0.13	0.14	0.02	.00	-0.02	-0.04	0.04
English	SRD SLOPE	0.09	-0.01	0.15	0.10	-0.05	0.12	0.10	-0.15	-0.15
	NRC SLOPE	0.16	0.17	0.21	0.18	-0.01	-0.06	-0.05	0.05	0.06
	ERC SLOPE	0.14	0.15	0.13	0.14	-0.14	-0.15	-0.12	0.02	-0.02
	NLC SLOPE	[illegible]	0.07	0.11	0.08	-0.02	0.05	0.02	-0.01	[illegible]
	ELC SLOPE	[illegible]	0.06	0.11	0.07	-0.08	.00	0.07	-0.08	-0.15

		Spanish								
		VDC I-INT	VDF I-INT	LDC I-INT	LSP I-INT	SRD I-INT	NRC I-INT	ERC I-INT	NLC I-INT	ELC I-INT
	VDC SLOPE	-0.08	0.02	-0.01	-0.23	-0.11	0.04	0.02	[illegible]	-0.05
	VDF SLOPE	-0.08	-0.05	-0.14	-0.06	-0.05	.00	[illegible]	[illegible]	-0.15
	LDC SLOPE	0.10	0.10	0.02	0.02	0.02	0.05	0.02	[illegible]	0.04
	LSP SLOPE	0.01	0.08	-0.09	-0.11	-0.07	-0.02	0.01	[illegible]	[illegible]
English	SRD SLOPE	-0.02	0.12	-0.04	-0.10	-0.07	-0.11	[illegible]	0.05	0.02
	NRC SLOPE	-0.07	-0.14	-0.12	-0.14	-0.01	0.08	0.05	-0.12	[illegible]
	ERC SLOPE	[illegible]	-0.05	-0.01	-0.03	0.18	0.15	[illegible]	[illegible]	-0.08
	NLC SLOPE	-0.03	[illegible]	-0.03	-0.05	.00	-0.04	[illegible]	0.13	[illegible]
	ELC SLOPE	-0.07	-0.03	-0.02	-0.07	-0.02	-0.01	[illegible]	[illegible]	[illegible]

Table 5

Interactive Reading Assessment System
Correlations Between English and Spanish
Growth Indices for the Bilingual Sample
(English student-intercepts vs Spanish growth indices)

		Spanish								
		VDC SLOPE	VDF SLOPE	LDC SLOPE	LSP SLOPE	SRD SLOPE	NRC SLOPE	ERC SLOPE	NLC SLOPE	ELC SLOPE
	VDC S-INT	-0.05	-0.01	0.01	-0.12	0.13	0.16	0.20	[illegible]	0.17
	VDF S-INT	-0.13	-0.17	-0.19	-0.27	-0.05	0.14	[illegible]	-0.02	[illegible]
	LDC S-INT	-0.04	-0.03	-0.01	-0.01	0.14	0.14	0.17	[illegible]	0.14
	LSP S-INT	[illegible]	-0.01	0.01	-0.06	0.16	0.18	0.20	0.17	[illegible]
English	SRD S-INT	0.17	0.02	0.05	-0.03	0.06	0.25	0.27	0.09	0.10
	NRC S-INT	[illegible]	0.09	0.03	-0.03	-0.01	-0.11	0.04	0.16	0.05
	ERC S-INT	0.02	0.05	-0.01	-0.02	-0.12	-0.08	0.06	0.02	-0.02
	NLC S-INT	-0.09	-0.01	-0.01	-0.10	-0.08	0.02	0.07	-0.15	-0.06
	ELC S-INT	0.01	0.02	-0.01	-0.09	-0.09	0.01	[illegible]	.00	[illegible]

		Spanish								
		DC S-INT	VDF S-INT	LDC S INT	LSP S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
	VDC S-INT	0.22	0.18	.10	0.26	0.13	-0.04	-0.11	0.17	-0.02
	VDF S-INT	0.08	0.14	0.11	0.17	0.05	-0.05	-0.06	-0.04	-0.07
	LDC S-INT	0.39	0.29	0.33	0.36	0.21	0.04	-0.11	0.17	[illegible]
	LSP S-INT	0.26	0.21	0.27	0.35	0.18	0.02	-0.12	[illegible]	[illegible]
English	SRD S-INT	0.16	0.17	0.14	0.24	[illegible]	-0.14	-0.19	[illegible]	0.13
	NRC S-INT	0.19	-0.15	0.05	0.13	0.16	[illegible]	0.05	-0.09	-0.02
	ERC S-INT	0.11	-0.02	0.14	0.17	0.12	0.26	[illegible]	[illegible]	0.07
	NLC S-INT	0.09	0.09	0.18	0.11	0.10	0.02	0.02	.02	0.17
	ELC S-INT	[illegible]	.11	0.11	0.18	0.23	[illegible]	-0.05	[illegible]	[illegible]

		Spanish								
		DC I-INT	VDF I-INT	LDC I-INT	LSP I-INT	SRD I-INT	NRC I-INT	ERC I-INT	NLC I-INT	ELC I-INT
	VDC S-INT	-0.16	-0.21	[illegible]	-0.12	-0.01	[illegible]	[illegible]	[illegible]	[illegible]
	VDF S-INT	[illegible]	[illegible]	-0.05	-0.12	-0.05	[illegible]	[illegible]	[illegible]	[illegible]
	LDC S-INT	[illegible]	-0.23	-0.29	-0.24	-0.11	[illegible]	[illegible]	[illegible]	[illegible]
	LSP S-INT	-0.33	-0.21	-0.29	-0.24	-0.07	[illegible]	[illegible]	[illegible]	[illegible]
English	SRD S-INT	-0.15	-0.22	-0.18	-0.14	-0.04	[illegible]	[illegible]	[illegible]	[illegible]
	NRC S-INT	-0.11	.18	[illegible]	[illegible]	-0.11	[illegible]	[illegible]	[illegible]	[illegible]
	ERC S-INT	-0.15	-0.04	-0.19	-0.22	-0.18	-0.20	[illegible]	[illegible]	[illegible]
	NLC S-INT	-0.11	-0.08	-0.11	-0.11	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	ELC S-INT	[illegible]	[illegible]	-0.15	-0.12	-0.06	[illegible]	0.14	[illegible]	[illegible]

538

Table [illegible]

Interactive Reading Assessment System:
Correlations Between English and Spanish
Growth Indices for the Bilingual Sample
English instruction-intercepts vs Spanish growth indices

		Spanish								
		VDC SLOPE	VDF SLOPE	LDC SLOPE	LSF SLOPE	SRD SLOPE	NRC SLOPE	ERC SLOPE	NLC SLOPE	ELC SLOPE
	VDC I-INT	0.04	0.14	0.02	0.12	-0.08	-0.09	-0.12	-0.18	-0.13
	VDF I-INT	0.15	.18	0.11	0.17	0.06	-0.06	-0.14	-0.05	-0.10
	LDC I-INT	-0.26	-0.04	0.08	0.01	-0.16	-0.15	-0.13	-0.12	[illegible]
	LSF I-INT	-0.12	0.14	0.04	0.15	-0.05	0.16	-0.23	-0.14	-0.13
English	SRD I-INT	-0.26	-0.01	-0.01	0.03	-0.08	-0.16	-0.20	-0.06	-0.08
	NRC I-INT	-0.05	-0.14	-0.04	.00	0.01	0.17	0.01	-0.05	-0.05
	ERC I-INT	-0.04	-0.07	-0.01	-0.03	0.13	0.07	-0.08	-0.02	-0.03
	NLC I-INT	0.07	0.12	0.07	0.11	0.02	-0.07	-0.15	0.08	0.02
	ELC I-INT	0.03	-0.02	0.04	0.11	0.03	-0.04	-0.15	0.04	0.07

		Spanish								
		VDC S-INT	VDF S-INT	LDC S-INT	LSF S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
	VDC I-INT	-0.15	-0.14	-0.10	-0.24	-0.11	0.01	0.05	-0.04	0.02
	VDF I-INT	-0.06	-0.16	-0.04	-0.10	-0.02	0.10	0.06	-0.01	0.08
	LDC I-INT	-0.41	-0.25	-0.43	-0.42	-0.25	-0.10	-0.04	-0.19	-0.15
	LSF I-INT	-0.30	-0.22	-0.31	-0.37	-0.23	-0.03	0.06	-0.08	[illegible]
English	SRD I-INT	-0.10	-0.16	-0.13	-0.18	-0.13	0.07	0.13	-0.14	[illegible]
	NRC I-INT	-0.05	0.02	-0.07	-0.12	-0.14	-0.11	-0.46	[illegible]	[illegible]
	ERC I-INT	[illegible]	-0.13	-0.14	-0.15	-0.36	-0.20	0.04	[illegible]	-0.15
	NLC I-INT	[illegible]	-0.09	-0.13	-0.13	-0.06	0.02	-0.01	[illegible]	[illegible]
	ELC I-INT	[illegible]	[illegible]	-0.14	-0.19	-0.20	-0.07	[illegible]	-0.04	[illegible]

		Spanish								
		VDC I-INT	VDF I-INT	LDC I-INT	LSF I-INT	SRD I-INT	NRC I-INT	ERC I-INT	NLC I-INT	ELC I-INT
	VDC I-INT	0.19	0.13	0.03	0.12	0.04	-0.04	-0.09	0.04	[illegible]
	VDF I-INT	0.09	0.13	.00	0.07	0.03	-0.07	-0.04	[illegible]	[illegible]
	LDC I-INT	0.22	0.22	0.44	0.28	0.18	0.05	0.12	0.17	[illegible]
	LSF I-INT	0.28	0.17	0.36	0.32	0.09	-0.05	[illegible]	-0.02	[illegible]
English	SRD I-INT	0.09	0.21	0.10	0.12	0.09	-0.09	[illegible]	[illegible]	0.07
	NRC I-INT	0.09	-0.09	0.10	0.10	0.11	[illegible]	0.05	[illegible]	[illegible]
	ERC I-INT	0.15	0.06	0.15	0.03	0.17	0.19	0.04	[illegible]	-0.02
	NLC I-INT	0.13	0.08	0.15	[illegible]	0.09	-0.12	[illegible]	[illegible]	[illegible]
	ELC I-INT	0.08	0.09	0.15	0.16	0.10	0.03	-0.09	[illegible]	.05

APPENDIX G

Between Language Correlations of IRAS Growth Indices and SFST Summary Measures

Table 1

Pre-reading and Reading Measures:
Correlations Between IRAS - English Growth Indices and
SFST - Spanish Summary Measures for the Kindergarten Bilingual Sample

IRAS - English (N=175)

Slope

		VDC SLOPE	VDF SLOPE	LDC SLOPE	LSP SLOPE	SRD SLOPE	NRC SLOPE	ERC SLOPE	NLC SLOPE	ELC SLOPE
	ALPHPR	0.06	-0.02	-0.03	0.08	0.11	0.06	0.15	-0.03	0.06
	WRDNAM	0.12	-0.11	0.11	-0.02	0.04	-0.01	0.08	0.09	0.11
	SNGLTM	0.02	-0.05	-0.01	0.18	-0.02	.00	0.04	.00	0.01
SFST	DBLLTM	0.06	-0.01	0.01	0.08	-0.07	0.01	0.09	-0.05	-0.06
Spanish	PSTRNG	0.08	0.07	0.01	0.13	0.01	0.07	0.11	0.04	0.06
Kinder	PSTRNF	0.14	0.07	0.06	0.15	0.09	0.15	0.25	-0.11	-0.03
	VCFDCL	-0.06	0.01	-0.12	-0.15	-0.27	-0.11	-0.09	-0.05	-0.03
	DEFNTN	0.02	0.14	-0.04	-0.10	-0.16	-0.05	-0.08	0.02	-0.01
	IMPFRB	-0.02	0.07	-0.06	-0.10	-0.16	-0.05	-0.04	-0.04	0.01

S-Intercept

		VDC S-INT	VDF S-INT	LDC S-INT	LSP S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
	ALPHPR	0.38	0.17	0.32	0.35	0.21	0.24	0.18	0.17	0.14
	WRDNAM	0.23	0.19	0.08	0.35	0.23	0.24	0.15	0.08	[illegible]
	SNGLTM	0.20	0.12	0.12	0.07	0.23	0.21	0.17	0.10	0.14
SFST	DBLLTM	0.25	0.13	0.11	0.16	0.33	0.20	0.11	0.16	[illegible]
Spanish	PSTRNG	0.12	-0.01	0.18	0.13	0.11	0.07	0.03	[illegible]	0.05
Kinder	PSTRNF	0.21	0.07	0.18	0.18	0.15	0.18	-0.02	0.21	0.20
	VCFDCL	0.17	0.06	0.16	0.24	0.24	0.05	0.01	-0.05	-0.04
	DEFNTN	0.29	-0.17	0.17	0.29	0.11	0.05	0.08	-0.16	-0.15
	IMPFRB	0.14	-0.10	0.18	0.30	0.14	0.08	0.08	0.05	0.01

I-Intercept

		VDC I-INT	VDF I-INT	LDC I-INT	LSP I-INT	SRD I-INT	NRC I-INT	ERC I-INT	NLC I-INT	ELC I-INT
	ALPHPR	-0.36	-0.17	-0.38	-0.16	-0.17	-0.26	-0.17	-0.19	-0.11
	WRDNAM	-0.08	-0.02	-0.12	-0.10	-0.07	-0.14	.00	-0.14	-0.17
	SNGLTM	-0.12	-0.02	-0.07	.00	-0.07	-0.14	-0.13	-0.18	-0.17
SFST	DBLLTM	-0.22	-0.05	-0.12	-0.06	-0.22	-0.12	-0.05	-0.24	-0.22
Spanish	PSTRNG	-0.16	0.05	-0.18	-0.10	-0.11	-0.09	-0.11	-0.16	-0.12
Kinder	PSTRNF	-0.19	-0.06	-0.10	-0.11	-0.13	-0.07	-0.04	-0.17	-0.14
	VCFDCL	-0.06	-0.01	-0.19	-0.19	-0.19	-0.12	-0.05	[illegible]	[illegible]
	DEFNTN	-0.03	-0.03	-0.20	-0.27	-0.09	-0.12	-0.07	[illegible]	[illegible]
	IMPFRB	-0.15	-0.05	-0.17	-0.32	-0.05	-0.16	-0.04	[illegible]	[illegible]

53[illegible]

Table 2

Pre-reading and Reading Measures:
Correlations Between IRAS English Growth Indices and
SFST - Spanish Summary Measures for the First-grade Bilingual Sample

IRAS - English (N=60)

Slope

		VDC SLOPE	VDF SLOPE	LDC SLOPE	LSP SLOPE	SRD SLOPE	NRC SLOPE	ERC SLOPE	NLC SLOPE	ELC SLOPE
	ALPHPR	0.16	0.25	-0.12	-0.20	0.06	0.01	0.06	0.09	0.04
	WRDNAM	-	-	-	-	-	-	-	-	-
	SNGLTM	0.01	0.01	-0.09	-0.16	-0.15	-0.26	-0.29	0.02	-0.11
SFST	DBLLTM	0.28	-0.08	0.30	-0.02	-0.14	-0.07	-0.07	-0.03	-0.23
Spanish	PSTRNG	0.10	0.14	0.02	-0.21	-0.17	0.13	0.13	0.19	0.11
First	PSTRNF	0.20	-0.06	0.14	0.03	-0.25	0.03	0.14	-0.03	-0.01
	VCFDCL	0.10	0.22	-0.08	-0.25	-0.19	0.12	0.10	0.17	0.02
	DEFNTN	-	-	-	-	-	-	-	-	-
	CMPRFB	-	-	-	-	-	-	-	-	-

S-Intercept

		VDC S-INT	VDF S-INT	LDC S-INT	LSP S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
	ALPHPR	0.01	-0.15	0.19	0.23	0.14	0.09	0.02	-0.11	0.04
	WRDNAM	-	-	-	-	-	-	-	-	-
	SNGLTM	0.23	0.14	0.13	0.18	0.25	0.35	0.36	0.22	0.20
SFST	DBLLTM	0.22	0.36	0.04	0.41	0.30	0.27	0.25	.31	0.29
Spanish	PSTRNG	-0.06	-0.14	.00	0.11	0.02	-0.11	-0.10	-0.12	-0.05
First	PSTRNF	0.15	0.25	0.10	0.29	0.24	0.05	-0.05	.21	0.10
	VCFDCL	0.06	-0.03	0.14	0.28	0.13	-0.08	-0.04	0.04	0.08
	DEFNTN	-	-	-	-	-	-	-	-	-
	CMPRFB	-	-	-	-	-	-	-	-	-

I-Intercept

		VDC I-INT	VDF I-INT	LDC I-INT	LSP I-INT	SRD I-INT	NRC I-INT	ERC I-INT	NLC I-INT	ELC I-INT
	ALPHPR	0.01	0.20	-0.18	-0.17	0.04	0.04	-0.03	0.14	0.12
	WRDNAM	-	-	-	-	-	-	-	-	-
	SNGLTM	-0.11	0.01	-0.09	-0.13	-0.27	-0.19	-0.24	-0.15	[illegible]
SFST	DBLLTM	-0.20	-0.20	-0.07	-0.30	-0.19	-0.44	-0.31	-0.19	[illegible]
Spanish	PSTRNG	0.03	-0.05	-0.07	-0.09	-0.07	0.17	0.10	[illegible]	[illegible]
First	PSTRNF	-0.14	-0.24	-0.11	-0.20	-0.14	-0.12	-0.07	-0.18	-0.07
	VCFDCL	-0.07	-0.04	-0.14	-0.24	-0.14	0.11	[illegible]	[illegible]	[illegible]
	DEFNTN	-	-	-	-	-	-	-	-	-
	CMPRFB	-	-	-	-	-	-	-	-	

510

Table 7

Pre-reading and Reading Measures:
Correlations Between IRAS - Spanish Growth Indices and
EFST - English Summary Measures for the Kindergarten Bilingual Sample

IRAS - Spanish (N=159)

Slope

		VDC SLOPE	VDF SLOPE	LDC SLOPE	LSP SLOPE	SRD SLOPE	NRC SLOPE	ERC SLOPE	NLC SLOPE	ELC SLOPE
	ALPHPR	-0.01	-0.01	-0.14	-0.15	0.12	0.12	0.11	0.04	[illegible]
	WRDNAM	-0.05	-0.02	-0.16	-0.19	0.11	0.05	0.09	-0.09	-0.04
	SNGLTM	-0.02	.00	-0.02	-0.03	0.08	0.06	0.03	0.02	-0.06
EFST	DBLLTM	-0.12	-0.13	-0.18	-0.25	-0.08	0.09	0.10	-0.11	0.06
English	PSTRNG	-0.18	-0.19	-0.19	-0.22	-0.11	-0.12	-0.05	-0.04	-0.01
Order	PSTRNF	-0.10	-0.18	-0.09	-0.11	-0.10	-0.13	-0.10	-0.15	-0.07
	VCFDCL	-0.36	-0.26	-0.26	-0.38	-0.22	-0.15	-0.12	-0.25	-0.09
	DEFNTN	-0.17	-0.12	-0.12	-0.13	0.00	.01	0.01	-0.05	0.11
	CMPPRB	-0.17	-0.18	-0.18	-0.22	-0.11	-0.06	-0.02	-0.17	0.01

S-Intercept

		VDC S-INT	VDF S-INT	LDC S-INT	LSP S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
	ALPHPR	0.09	0.09	0.20	0.24	0.05	-0.12	-0.11	-0.01	-0.07
	WRDNAM	0.14	0.10	0.15	0.20	0.07	-0.03	-0.09	0.03	0.01
	SNGLTM	0.08	0.13	0.11	0.12	0.01	-0.12	-0.08	0.09	0.15
EFST	DBLLTM	0.09	0.13	0.14	0.21	0.12	-0.11	-0.16	0.19	0.07
English	PSTRNG	0.12	0.09	0.17	0.15	0.15	0.12	0.03	-0.05	-0.02
Order	PSTRNF	0.06	0.16	0.10	0.11	0.14	0.11	0.06	0.08	0.08
	VCFDCL	0.09	0.10	0.09	0.15	0.09	0.08	[illegible]	-0.05	-0.09
	DEFNTN	0.15	0.22	0.11	0.12	0.09	-0.01	[illegible]	0.11	-0.02
	CMPPRB	0.05	0.15	0.08	0.10	0.13	0.05	-0.04	0.21	0.10

I-Intercept

		VDC I-INT	VDF I-INT	LDC I-INT	LSP I-INT	SRD I-INT	NRC I-INT	ERC I-INT	NLC I-INT	ELC I-INT
	ALPHPR	-0.08	.00	-0.07	-0.13	-0.01	0.10	0.07	[illegible]	[illegible]
	WRDNAM	-0.01	0.01	0.01	-0.03	-0.12	.00	[illegible]	[illegible]	[illegible]
	SNGLTM	0.06	-0.07	-0.09	-0.01	0.03	0.08	[illegible]	[illegible]	-0.06
EFST	DBLLTM	-0.05	-0.13	-0.12	-0.11	-0.10	0.06	[illegible]	-0.17	[illegible]
English	PSTRNG	-0.08	0.04	-0.02	-0.10	-0.15	-0.13	-0.06	0.04	[illegible]
Order	PSTRNF	-0.04	-0.01	-0.01	-0.09	-0.14	-0.15	-0.09	[illegible]	[illegible]
	VCFDCL	-0.04	0.01	0.07	-0.12	-0.16	-0.15	-0.08	-0.01	[illegible]
	DEFNTN	-0.14	-0.10	0.13	0.05	-0.04	-0.05	-0.14	[illegible]	[illegible]
	CMPPRB	-0.08	-0.13	0.09	0.14	-0.10	-0.17	0.04	[illegible]	[illegible]

541

Table 4

Pre-reading and Reading Measures:
Correlations Between IRAS - Spanish Growth Indices and
SFST - English Summary Measures for the First-grade Bilingual Sample

IRAS - Spanish (N=49)

Slope

		VDC SLOPE	VDF SLOPE	LDC SLOPE	LSP SLOPE	SRD SLOPE	NRC SLOPE	ERC SLOPE	NLC SLOPE	ELC SLOPE
	ALPHPR	0.16	0.06	0.26	-0.02	0.15	0.18	0.13	-0.15	0.14
	WRDNAM	-	-	-	-	-	-	-	-	-
	SNGLTM	0.12	-0.14	0.04	0.01	0.01	0.06	0.19	0.11	-0.06
SFST	DBLLTM	0.23	0.29	-0.13	-0.02	-0.07	0.02	0.22	0.34	0.19
English	PSTRNG	0.16	-0.04	0.09	0.05	0.27	0.23	0.17	-0.02	0.06
First	PSTRNF	0.11	0.26	-0.07	-0.07	0.09	0.10	0.14	0.23	0.35
	VCFDCL	0.06	-0.05	0.07	-0.15	0.17	0.20	0.30	0.06	0.17
	DEFNTN	-	-	-	-	-	-	-	-	-
	CMPPRB	-	-	-	-	-	-	-	-	-

S-Intercept

		VDC S-INT	VDF S-INT	LDC S-INT	LSP S-INT	SRD S-INT	NRC S-INT	ERC S-INT	NLC S-INT	ELC S-INT
	ALPHPR	0.11	-0.02	0.10	0.38	0.06	-0.05	-0.04	0.03	-0.12
	WRDNAM	-	-	-	-	-	-	-	-	-
	SNGLTM	0.09	0.34	0.14	0.20	0.18	0.05	-0.12	0.17	0.14
SFST	DBLLTM	0.17	0.05	0.44	0.36	0.28	0.13	-0.09	-0.08	-0.03
English	PSTRNG	-0.17	0.02	0.06	0.09	-0.14	-0.20	-0.17	-0.01	-0.05
First	PSTRNF	0.03	-0.02	0.24	0.24	0.08	-0.02	-0.05	-0.02	-0.20
	VCFDCL	0.05	0.11	0.09	0.26	0.00	-0.08	-0.23	0.12	-0.03
	DEFNTN	-	-	-	-	-	-	-	-	-
	CMPPRB	-	-	-	-	-	-	-	-	-

I-Intercept

		VDC I-INT	VDF I-INT	LDC I-INT	LSP I-INT	SRD I-INT	NRC I-INT	ERC I-INT	NLC I-INT	ELC I-INT
	ALPHPR	-0.08	-0.14	-0.17	-0.43	-0.14	0.14	0.14	0.01	0.19
	WRDNAM	-	-	-	-	-	-	-	-	-
	SNGLTM	-0.19	-0.22	-0.08	-0.22	-0.29	-0.11	0.11	0.01	0.02
SFST	DBLLTM	-0.27	-0.04	-0.41	-0.40	-0.36	-0.17	0.02	[illegible]	0.07
English	PSTRNG	0.08	0.02	-0.05	-0.16	-0.01	0.25	0.16	0.02	0.02
First	PSTRNF	-0.04	-0.01	-0.20	-0.25	-0.20	0.08	0.08	-0.11	0.10
	VCFDCL	-0.10	-0.30	-0.15	-0.28	-0.07	0.11	0.26	-0.02	0.01
	DEFNTN	-	-	-	-	-	-	-	-	-
	CMPPRB	-	-	-	-	-	-	-	-	-

Final Report

TEACHING READING TO BILINGUAL CHILDREN STUDY

Volume 6
Instruction

Wesley A. Hoover
Robert C. Calfee
Betty J. Mace-Matluck

Document BRS-84-R. 1-VI

Preston C. Kronkosky
Executive Director

Southwest Educational Development Laboratory
211 East Seventh Street
Austin, Texas
(512) 476-6861

November 1984

Final Report

TEACHING READING TO BILINGUAL CHILDREN STUDY

Volume 6
Instruction

Wesley A. Hoover, Robert C. Calfee
Betty J. Mace-Matluck

Document BRS-84-R. 1-VI

Preston C. Kronkosky, Executive Director
Southwest Educational Development Laboratory
Austin, Texas

November 1984

There were many individuals and institutions who contributed to this research effort. We wish to express our sincer gratitude to the parents, students, and school personnel who provided the necessary data from which this study is derived.

In addition, several other individuals made valuable contributions to the study, for which we are indebted: Robert C. Calfee, Sylvia C. Peña, and Blanca de Alvarez.

And finally, we wish to thank the local data collectors at the school sites, many of whom remained with the study throughout its duration: Ramiro Barrera, Beatrice Cantú, Irene Cavazos, Carolyn Cruz, María de Obregón, Gloria de Torres, Gigi Galván, Olga Hernández, Irene Méndez, Guadalupe Treviño, Rosalinda Villarrando, and Gloria Villarreal. Their patience, dedication, and hard work helped make this study a reality.

Betty J. Mace-Matluck
Wesley A. Hoover

TABLE OF CONTENTS

LIST OF TABLES

517

Table

Page

Table Page

Table Page

Table

Page

LIST OF FIGURES

Figure

551

PREFACE

In June 1978 the National Institute of Education (NIE) funded the Southwest Educational Development Laboratory (SEDL) to conduct a longitudinal study on the Teaching of Reading to Bilingual Children. Educators and policymakers alike have long recognized that the ability to read is essential for success in school, in work, and in life; yet many children from second-language backgrounds have trouble learning to read in schools today. The majority of these youngsters are from Spanish-language backgrounds and from low income families. Special programs designed to meet the needs of these children are provided in schools, but there is limited research evidence to guide the development, evaluation, and implementation of these programs. This study is intended to provide information that will result in greater insights into what constitutes a favorable learning environment for children from Spanish-language backgrounds, what instructional sequences and events promote successful and efficient learning of literacy skills, and what the language and literacy outcomes of current schooling practices are for a large sample of these youngsters.

The study was conducted during the years of 1978 through 1984. It is a comprehensive longitudinal investigation of the development of reading skills from kindergarten through fourth grade for a representative sample of more than 350 children from bilingual backgrounds, and for smaller samples of children who, on entry into school, were monolingual in English or Spanish. In this "natural variation" study, teaching and learning were carefully documented in field settings at the several sites.

The goals of the study were to (a) describe variations in both English and Spanish language ability of students living in bilingual communities, (b) document prevailing practices in reading instruction for bilingual students, and c) investigate the relations between the instructional program and student achievement for students with differing entry profiles.

Description of the Study

Surveys of the general and school populations reveal an increase in the number of students whose language resources are not an ideal match to the language of the school. An important question for educational practice and policy centers around the school's responsibilities in this situation. Bilingual programs, English-as-a-Second-Language classes, classroom aides, and "sink-or-swim" approaches can all be found in practice today. From limited evidence now available, none of these techniques has emerged as the one best system.

Hispanics make up the largest and fastest growing school-age population today. The demographics for some states show that over the next decade they may constitute as much as a third to a half of the population. In the state of Texas at present approximately one third of the school children are from Hispanic backgrounds (approaching one

million). They are found in virtually ever school district in the state. Many of the school districts in the southern portion of the state serve school populations of which 75% to 99% of the children are from Spanish-speaking backgrounds and, on entry into school, are often limited in their ability to speak English and to profit from instruction in that language. This population is not restricted to the border areas, however. Large urban centers in the state report as much as 20% of their school population from Hispanic backgrounds, with a concentration of some 80% to 90% in certain of their schools.

It is well documented that, in general, children from Spanish-speaking backgrounds, for whatever reason, often encounter difficulty in our nation's schools; they do more poor y on standardized tests than does the general school population, and their dropout rate is high. Bilingual education, in which students are given instruction partially through the home language until they have attained sufficient proficiency in English to benefit from English-medium instruction, has been the principal approach recommended by the Office for Civil Rights to ensure access to equal educational opportunity for these children. Although many individual programs have had considerable success in improving the academic performance of language-minority students, it has not been demonstrated that these programs generally are reducing inequality of educational opportunity on the large scale that was envisioned.

Growth in reading comes about for most youngsters through formal classroom instruction. Understanding the development of reading, and knowledge of the critical variables that determine success or failure, depends on a careful examination of the instructional program -- not just the label over the classroom door, but the program as actually implemented by the classroom teacher.

Educators have raised several issues about the most effective way to help bilingual children become proficient readers of English. These include (a) valid assessment of the student's ability in the languages of the home and of the school, (b) the optimal balance of formal instruction in both languages, (c) the most effective transfer from one language to the other, and (d) bilingual support within the classroom environment. A major thesis of the <u>Teaching Reading to Bilingual Children</u> study is that addressing these issues (and others) requires a comprehensive and ecologically-valid investigation of the linkage between the child's language and the language of instruction.

Design of the Study

To achieve the objectives of the study, considerable attention was given to the selection of schools, teachers and students, to the instruments for assessing language and reading achievement, and to the methods for evaluating the classroom instruction. Each of these topics is discussed briefly below.

Schools, Classes and Teachers

Twenty schools and 200 teachers from six school districts participated in the study. Included are variations in the nature of the reading program (a range from phonics-oriented to meaning-based), classroom organization (some self-contained, others team-taught), and grade structure (the range of grades in the individual school and the extent of cross-grading both vary). The schools differed in size, SES, urbanicity, locale, and makeup of the student body (from medium to high concentration of bilingual students).

Student Cohorts

The study was undertaken in four cohorts or "waves" of students. Three of the cohorts consisted entirely, or in large part, of bilingual students. The first cohort was small (N=40) and of limited generality; the second was somewhat larger (N=80) and covered a slightly broader array of contexts. The third cohort which was both larger (N=200) and broader in its generality, incorporated a number of procedural improvements based on previous experience in the study and included a monolingual English-speaking sample. The fourth cohort consisted of a relatively small sample (N=60) of monolingual Spanish-speaking students.

All of the bilingual sites were from the state of Texas, as were the monolingual English-speaking students. The monolingual Spanish-speaking students were from one site in Northern Mexico.

The original design of the study called for each student to be assessed and observed from entry to kindergarten through exit from third grade. By covering the full range of the primary years, we would be able to examine the transition from "learning to read" through "reading to learn." For students in programs where the initial stages of reading were in Spanish, we also considered it important to determine the transition to competence in English reading.

The original design was in fact implemented for the first two cohorts; some of the students were tracked from first through fourth grade, but most followed the intended design. Due to limited funding in the later stages of the study the last two cohorts could not be followed for the full four years that were originally intended. The bilingual and monolingual English samples from the Texas sites were observed from kindergarten through second grade, and the monolingual Spanish samples from the site in Northern Mexico were observed from first through third grade (the program did not provide a kindergarten).

The monolingual samples were incorporated in the design to aid in validating the instruments for student assessment. Both the English and Spanish cohorts are small and not selected to be fully representative of monolingual populations. Data from these samples will be presented in Volume 3, as part of the discussion on the adequacy of the instruments for measuring growth. The study was designed to study the course of reading in bilingual students, not as a basis for comparing these students with monolingual youngsters. Accordingly, comparisons

between the various samples will not be made in this report, nor do we recommend that others attempt such comparisons.

Language Assessment

Several types of data were collected for each student on English and Spanish proficiency. Each year, early in the Fall and again in the Winter and Spring, teachers rated their students' language skills. Oral language proficiency tests were administered in the Fall of each year. Finally, audiotaped speech samples were obtained monthly on a rotating schedule in three settings: in the classroom, on the playground, and in the home.

Reading Assessment

Several instruments were used to measure reading achievement. Standardized test scores (mostly English), were collected yearly. More detailed information was obtained from a battery of individually-administered "performance based tests" in both English and Spanish. In kindergarten, the Stanford Foundation Skills Test was employed to measure the child's pre-reading skills. From the end of first grade on, the Interactive Reading Assessment System was administered during the Spring of each school year. This instrument provides independent measures of the student's skills in decoding, word meaning, fluency in oral reading, and comprehension. Finally, informal reading inventories were administered throughout the school year.

Classroom Observations and Teacher Interviews

Project staff conducted monthly observations of the reading instruction in each classroom and interviewed the teachers quarterly about their instructional plans. The observation instrument documented staffing patterns, grouping and organization, time allocation, the language of instruction, the character of instruction, the materials and procedures used, and the response of the students. The interviews focused on the teacher's general instructional objectives, as well as the objectives for individual target students. Taken together, these two instruments yield a rich characterization of the classroom environment for the target students.

Student Entry Variables, Classroom Factors, and Reading Achievement

The primary goals of the analyses were to identify the general relationships that characterize variation in these factors and to look for underlying regularities that are associated with success and failure, both in the early stage of reading instruction and in the year-to-year variations.

Docum

This report is one of a series of eight documents contained in the Final Report submitted to the National Institute of Education. A com-

plete list of these documents is provided on the inside of the cover of this report.

The study was a collaborative effort among a number of individuals and institutions. All members of the research team contributed to the thinking, planning, and writing of this series of documents, however, the individual whose name appears first in the list of authors was responsible for preparing the particular document.

Betty J. Mace-Matluck
Wesley A. Hoover
Co-Principal Investigators

Austin, Texas
November 30, 1984

INTRODUCTION

In the previous volume, the reading performance of the bilingual sample was discussed, and substantial individual differences in the patterns of growth in both English and Spanish reading were evident. One of the main thrusts of this study is how such differences in growth can be explained, and in particular, the role played by instruction. As such, the study invested large resources in acquiring information about the specific instruction the target sample received over the early elementary grades in which their progress in reading was being tracked. In this volume, these sources of information, and the data obtained from them, are described. In the next volume, the link between these individual instructional indices and student deviations in reading growth patterns will be discussed.

A coordinated system employing classroom observations and teacher interviews was used in the study to (a) obtain detailed characterizations of the classroom instruction experienced by each of the students in the study (b) document the teachers' general instructional objectives, as well as those for individual students, (c) describe the nature of the instructional program, both at the school and classroom level, and (d) gather information about the teachers' background, training, and language skills. The primary data sources employed for describing individual student instruction are the Reading and Mathematics Observation System (RAMOS) and the Teacher Checklists (Checklists). These are the focus of this report and are described in detail below. The remaining instruments, of which there are three, are ancillary information sources that have utility for cross-validation of the primary sources. These include (a) the Bilingual Classroom Questionnaire (BCQ), and its successor, the Inventory of Bilingual Instruction (IBI), and (b) the Survey of Teachers' Background and Language Skills (STBLS). Both the BCQ and IBI are based on teacher interviews about daily teaching schedules (e.g., subjects taught, languages employed, length of instructional period, language background of students), and thus provide information relevant to instructional language usage across each of the language arts and content areas taught in the school. The STBLS, again based on teacher interview, provides information about each individual teachers' educational background, training, and self-assessed skill in using both English and Spanish in various settings ranging from the very informal to the very formal. These three instruments (BCQ, IBI, and STBLS) are not treated in this document; however, a copy of each of the instruments is provided in Appendices A and B, for the BCQ/IBI and the STBLS, respectively.

For the two instruments that are the focus of this report (RAMOS and Checklists), the specific structure of each is presented below, along with the procedures employed for deriving summary indices. This is followed by a discussion of the descriptive statistics for these derived summary indices of instruction for the bilingual target sample. Finally, a description of the methods employed in deriving an index of the primary language of the reading program over the primary grades (years of Spanish reading instruction) is presented.

READING AND MATHEMATICS OBSERVATION SYSTEM

The primary source of instructional data was obtained from regular classroom observations. For each target student, these observations were conducted during the scheduled reading period of those teachers with primary responsibility for providing reading instruction to the target student. In this section, the structure of the observation instrument is first described, followed by discussions of the data collection and processing procedures, the derivation of summary indices, and descriptive statistics on the aggregated measures for the bilingual sample. These are followed by descriptions of further data reduction procedures which yielded factor scores, and a presentation of their associated descriptive statistics and inter-correlations.

Instrument Description

The Reading and Mathematics Observation System (RAMOS) was developed by Robert and Kathryn Calfee (Calfee & Calfee, 1974, 1976) as a method of recording instructional events occurring within elementary school classrooms. Primarily designed to document the presentation of reading, language arts and mathematics lessons, the system is flexible, providing for easy modification. For example, the set of codes used to characterize reading activities can be expanded to suit particular needs, as was done in the Reading Diary Study (Piontkowski, 1981) to encompass the broader range of skills and materials more frequently found at the primary level.

The RAMOS system provides real-time documentation of classroom instruction, unlike the time-sampling methods used in many other systems. Using a special form, a trained observer begins an observation by recording the current status of the instruction delivered as applicable under each of the RAMOS categories, thus providing a baseline account of the initial structure and content of the instruction being offered. From then until the end of the observation, any time there is a change in skill or activity in one or more of the specified instructional categories, this change is identified and recorded along with the time at which the change occurs. Such documentation allows an assessment of both the quality (as represented by the particular code) and quantity (as represented by the number of minutes a particular code is employed) of effort spent by the teacher and students under each of the several categories of RAMOS.

Information is recorded on an Event Form using computer-compatible mnemonic codes contained on a Master Code Sheet. The Event Form is designed to answer some of the following questions about classroom activities:

What is the classroom organization?

> What is the structure of the groups of students? How are they organized?

Where are the target students (specific children selected to be observed as individuals)? In what groups?

Where are the instructors?

Who is doing the instructing, and how?

What is the classification of the instructor (teacher, aide, volunteer, tutor, etc.)?

What role is the instructor playing (direct instruction, discipline, classroom management, assessment, etc.)?

What is the content of instruction?

What subject matter is being taught?

If reading is being taught, what sk ; and activities are involved?

What materials are being used?

What is the response of the students to instruction?

What are the tasks required or expected of the students?

What is the judged level of attention to instruction? What is the relative adequacy of performance? What amount of social interaction is there between students?

In what ways, if any, are the activities of target students different from those of the group?

With these general questions in mind, the RAMOS system was adapted to meet the specific needs of the Teaching Reading to Bilingual Children Study. In this version (Calfee & Calfee, 1978), the basic format and content of the system were retained. Some categories were removed from the Event Form, modified to obtain information of a more global nature, and included in the Rating Summary Sheet (not discussed in this report), which was completed by the observer immediately following each observation. Other changes occurred in the content section of the Event Form where the reading codes were modified to permit documentation specific to the SEDL research (changes mainly concerned with the focus of instruction an materials, the teaching strategies employed, and the language us d during instructional events) In Appendix C, the RAMOS/SEDL Event Form, Rating Summary Sheet, Master Code Sheet, Definitions of Codes, and Definitions of Scales are presented; each (except for the Rating Summary Sheet) is discussed fu her below.

The Event Form (see Appendix C) is a record in coded form of what occurs in the classroom. The top of the page is used for basic information: place, date, observer, names of any s d materials used, and the initial identification and assignment of the target students and the instructors. The instructional categories contained on the Event Form for capturing the subsequent sequence of classroom activities are summarized below (refer to the Master Code Sheet in Appendix C for the specific codes allowed under each category):

Time: the actual clock time when the activities coded for a given group begin (to the nearest minute on a 24 hour clock).

Target Student: the unique identification number of a given target student, entered when the coded information line pertains only to that particular target student and not to the entire group to which the target student is currently assigned.

Status: the status of the entered line of information, used to indicate whether (a) an instantiated group is (or is no longer) being observed, (b) a target student is involved in a different instructional activity than the one the group is generally undertaking (or is returned to doing the same activity as the other students in the assigned group), or (c) a given activity is deemed to be important, but lasts less than thirty seconds (a momentary activity).

Group ID: the unique identification number of the group being monitored.

Group Number: the number of students assigned to the group under observation.

Instructor ID: the unique identification number of the instructor associated with the group being observed.

Instructor Classification: the classification of the instructor associated with the group being observed (e.g., teacher, aide, volunteer).

Instructor Role: the role the instructor is playing in delivering the instruction to the group being observed.

Subject Matter: the nominal curriculum in use by the group under observation.

Instructional Focus: the instructional emphases and strategies of the instruction undertaken by the group being observed (under six skill headings of general, grammar, vocabulary, decoding, comprehension, and interpretation skills).

Technique: the manner in which skills of visual or auditory pattern recognition are presented by the instructor associated with the group under observation (as parts-to-whole or whole-to-parts).

Language of Instruction: the language employed by the instructor in the delivery of instruction to the group under observation.

Primary Materials: the materials being used by the group under observation which are the central focus of the instructional event.

Ancillary Materials: the supplemental materials used by the group under observation which are secondary to the primary materials in their function.

Activity/Task: the type of activity chosen as a vehicle to convey the instructional content to the group being observed (under nine activity headings of lecture, discussion, independent work, questions/answers, recitation, audio-visual, transitional, other reading-related activities, and non-reading activities).

Attention (Collection Years 1-2): for the group being observed, the observer's rating of the overall attention of the students to the assigned task.

Number of Nonengaged Students (Collection Years 3-5): the number of students in the observed group who are not attending to the task assigned.

Productivity: for the group being observed, the observer's rating of the productivity of those students engaged in the assigned task.

Noise/Social Interaction: for the group being observed, the observer's rating of the noise level of the group (relative to the activity/task being performed).

As noted above, the category of Attention was only employed in the first two years of data collection; in the remaining three years, it was replaced with the more specific index, Number of Nonengaged Students -- this was the only major change made in the RAMOS instrument during the five years of data collection.

An example taken from a completed protocol will be used to demonstrate the method this system employs in characterizing instruction in the current study. Figure 1 (presented on the next two pages) displays a completed sample Event Form taken from a Site 5, second-grade classroom during the last year of data collection. This example will be the basis of further examples used in describing the derivation of instructional summary indices, so understanding it will facilitate the understanding of subsequent discussions of procedures.

In the top panels of the Event Form, four target students are identified; the first three are assigned to group 1 (containing 19 students with no associated instructor), and the fourth to group 2 (consisting of six students with the teacher). Interpreting the first line of the protocol (from left to right), group 1 is formed at

Group No.	ID	Unit	
1	19	-	-
2	6	1	EV

Target Students				
1	AG	1	-	5099
2	CG	1	-	5109
3	RG	1	-	5120
4	AS	2	1	5289
5				
6				
7				
8				
9				
10				

Reading Curriculum Materials
MAGIC TIMES
MACMILLAN
GROUP 3: p. 210-215
GROUP 4: p. 196-201

Other Instructors

[redacted]
[redacted]
Date 012483
[redacted]
Time in 0830
Time out 0920

Time 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
0	8	3	0		N	1	1	9	0	X	X	L	S	W	S	X	E	C	B	P	P	W	C	M	0	0	M	L						
0	8	3	0		N	2	0	6	1	T	N	L	D	W	W	X	E	W	G	Q	T	R	R	R	0	0	M	L						
0	8	3	5			2							G	C	W																			
0	8	3	7			2			0	X	X		S	W	S			C	B	P	P	W	C	M										
0	8	3	7		N	3	0	5	1	T	N	R	S	S	L	P	E	Q	T	X	X	Q	S	A	0	0	H	L						
0	8	3	7	2		3																												
0	8	3	7	3		3																												
0	8	3	7			1	1	4																										
0	8	3	9			3					R		X	X	X			X	X			L	L	S			X							
0	8	4	1			3					N											Q	S	A			M							
0	8	4	2			1	1	1																										
0	0	0	5	Children left for ESL language Lab																														
0	8	4	2			3					F		S	O	R	X		B	R			R	R	A			M							
0	8	4	5			3					M		C	L	F					Q	T	Q	S	A										
0	8	4	6			3					F		S	O	R					X	X	R	R	A										
0	8	5	3			3					N		C	L	F					Q	T	Q	S	A			L							
0	8	5	6			3					F		S	O	R					X	X	R	R	A			M							
0	8	5	8			3			0	X	X	L	S	W	S			C	B	P	P	W	C	M										
0	8	5	8		N	4	0	4	1	T	N	R	G	C	N	P	E	C	H	X	X	R	R	R	0	0	M	L						
0	8	5	8	1		4																												

RAMOS/SEDL Event Form
Copyright February 1974 (Revised 10/75; 10/78)
R. Calfee and K. H. Calfee, Stanford, California.
Adapted for use in SEDL research 9/1/78 (Revised 8/80). Not to be reproduced or used without express permission of the authors.

Figure 1. Reading and Mathematics Observation System: Sample raw protocol from a Site 5 second grade.

562

BEST COPY AVAILABLE

READING CURRICULUM MATERIALS

OTHER INSTRUCTORS

[redacted]

[redacted]

DATE Ø12483

[redacted]

TIME IN

TIME OUT

GROUP | INSTRUCTION | CONTENT (READING) | ACTIVITY/TASK | RESPONSE

TIME (1–4)	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
ØØØ5	Group 4 reads sentences using both the contraction and the two words that make up the contraction																								
Ø958			1	Ø	7																				
Ø9ØØ	1	D	4																R	R	A				
Ø9Ø1	1	R	4																						
Ø9Ø6			4							A	C	S			Q	T			Q	S	A				
Ø9Ø6		M	3			1		C																	
Ø9Ø7			4					F		S	O	R	X		B	R			R	R	A				
Ø91Ø			4					N		C	L	F			Q	T			Q	S	A				
Ø911			1	1	Ø																				
ØØØ5	Children return from ESL lab																								
Ø913			2	Ø	4																				
ØØ45	Children leave for ...																								
Ø917			4			Ø	X	X	L	S	W	S			C	B	P	P	W	C	M				
Ø917			1			1		F																	
Ø96Ø			Ø																						
ØØ99																									

RAMOS/SEDL Event Form

8:30; it consists of 19 children with no instructor present ('0' in column 10), and thus, instructor classification and role codes do not apply ('X' in columns 11 and 12). The subject is language arts ('L' in column 13); the instructional focus is sentence writing ('SWS' in columns 14-16), and thus, technique does not apply ('X' in column 17). The language of instruction is English ('E' in column 18), with the chalkboard constituting the primary material, and paper and pencil used as supplemental materials ('C'' and 'P' in columns 19-20 and 21-22, respectively). The activity/task is independent work consisting of copying material ('WCM' in columns 23-25), and each of the students is engaged ('00' in columns 26-27). The productivity of the group is rated as medium with the noise level low ('M' and 'L' in columns 28 and 29, respectively). In sum, the classroom event coded for group 1 is one where students are working independently at their desks, copying English sentences from the chalkboard with paper and pencils. By scanning down the Group ID category (column 7) and locating the references to group 1, one can see that this characterization continues throughout the 51 minutes of observation with only a few modifications: at 8:37, five students are taken from group 1 to form group 3; at 8:42, three additional students leave the group for an ESL language laboratory (note that comment lines are coded with '0005' under the time category); at 8:58, four more students leave the group to form group 4; at 9:11, the three students who left earlier for the ESL language laboratory return; and at 9:17, the teacher begins facilitating the seatwork of the group until the observation is ended at 9:20.

Group 2 is also formed at 8:30, containing six students and the teacher, who is providing direct instruction ('N' in column 12). The subject matter is again language arts, but the instructional focus is on whole word recognition ('DWW' in columns 14-16). English is the language of instruction; the primary materials are word cards, with secondary materials being questions from the teacher ('WC' and 'QT' in columns 19-20 and 21-22). The activity/task is reading and responding ('RRR' in columns 23-25). All students are engaged, productivity is moderate, and the noise level is low. Overall, the students are engaged in teacher directed instruction aimed at sight-word recognition using word cards. This activity continues for five minutes when the focus switches from whole word recognition to compound words -- one or more compound words have appeared in the word cards, and for the next two minutes the teacher provides instruction on their structure. At 8:37, group 2 is assigned the same task as group 1 (i.e., copying material from the chalkboard), while the teacher begins work with a third group of five children formed from a subset of the students in group 1.

Given this basis for interpretation, the reader is left to follow the instruction documented for groups 3 and 4. However, the use of four important conventions require explanation in the example presented in Figure 1. First, note that at 8:37, group 3 is formed, and in the two lines following the formation of this group, the assignments of target students 2 and 3 to this group are made -- employing this procedure of tracking successive group assignments, individual target students may be tracked throughout the observation, thus providing information on which to build individual instructional profiles.

The second event worth special notice occurs at 9:00. Here target student 4 is observed doing something different from the assigned group ('D' in column 6), and the difference indicated appears in the activity/task category. While group 4 is generally involved in reading and responding, target student 4 is asked to read aloud ('RRA' in columns 23-25) for a minute; he is returned at 9:01 to the activity/task of the group. Using this procedure allows the documentation of instruction for an individual student even when that instruction is different from the group to which the student is nominally assigned.

The third point of special interest concerns the handling of a fairly frequent event, that of momentary control. In Figure 1, at 9:06 such a momentary activity is noted ('M' in column 6). The only other entries for the line concern the instructor, and here, the teacher is observed to be acting in a control mode ('C' in column 12) over group 3. Prior to this time, the teacher was providing instruction to group 4, but some event within group 3, working independently on the assigned copying task, required the teacher's attention for a short period of time (less than 30 seconds). Had the control required more time, then the teacher would have been assigned to group 3 and removed from group 4. By using this technique, an assessment of the number of interruptions requiring an instructor's management within an instructional event can be made.

The final point of special interest concerns the last line of the observation which is always '0099' entered in the first four columns. This is simply the special code recognized by the analysis program to mark the end of the observation.

Before turning to the procedures used to summarize the RAMOS protocols, a discussion of the collection and initial processing of the raw protocols will be given.

Data Collection and Processing

As discussed in Volume 2 (page 43) of this series of reports, the study was quite fortunate in acquiring the skills of an exceptionally able group of data collectors at each of the six sites -- to our delight, most remained with the study for its duration. Given the complexity of the observation system, substantial training is required, but the teaching experience each of the data collectors possessed was quite useful in facilitating their acquisition of its principles. In the late summer of the year of a site's initial entry into the study, extensive training on the RAMOS (lasting about five days) was provided by personnel expert in its use. Supplemental on-site training followed in mid-Fall and late Spring (each of about two days duration); in subsequent years, three training sessions (of somewhat shorter duration) were held at about these same times. In the fourth year of data collection, all site personnel were brought to SEDL for two days of training in an effort to ensure uniformity in classroom coding.

The data collection schedule (see Volume 2) called for monthly classroom observations (October through April) of the reading instruction given each target student by the teacher with primary responsibility for such instruction. The schedule was difficult to follow given the other duties the study assigned to the data collectors, and the spreading of target students after their initial selection to all available classrooms in subsequent grade levels. As a result, only five to six of the scheduled seven observations were generally obtained each year.

Completed protocols were received at SEDL on a monthly basis, and in-house staff checked their coding for completeness and accuracy. Continual communication between SEDL staff and field observers was essential, and allowed uniform handling of special cases as they were encountered. Once checked, the raw protocols were then entered into computer files, frequency analyses conducted, and identified coding/entry errors corrected. As detailed below, the raw alphanumeric codes were then transformed to ordinal numeric codes based on scaling routines, and summary indices of the instructional dimensions of interest for each target student in each observation were generated. These files were then segregated into English and Spanish reading observation summaries. At the end of the data collection phase of the study, these language-specific collection year summary files were then merged to create language-specific instructional year summary files (i.e., English and Spanish summary files organized by instructional year, kindergarten through fourth grade).

Segregation of English and Spanish Protocols

As mentioned above, once entered into computer files, individual target student summaries (i.e., summaries based on individual observations) were segregated into English reading instruction and Spanish reading instruction observations. These distinctions were made mainly on the basis of the primary language of the materials used in the instructional groups each target student was assigned (information about the language of the materials was contained at the top of each Event Form). Similar information about the language of materials as indicated on any Checklists obtained during the same period of time (discussed in the next major section) were also reviewed, and any discrepancies between these information sources were checked with the field observer (or relevant teacher) to determine whether a given student at a given time was primarily enrolled in an English, Spanish, or dual language reading program.

Once these files were created, the individual student summaries were checked to determine the amount of time on which each summary was based (i.e., the number of minutes the student was actually observed in each summarized protocol). Since some summaries were found to be based on little actual observation time (e.g., students leaving the classroom to attend language labs held in other rooms), all summaries based on 10 or fewer minutes were removed from the files and did not enter into any subsequent analyses. Of the remaining protocols, the average observed time was about 40 minutes (see the tables presented below which provide these statistics for each instructional year). We now turn to a discussion of the procedures used to summarize the raw RAMOS English and Spanish reading protocols.

Derivation of Summary Measures

Instructional dimensions were quantified by constructing scales from the raw codes under each of the RAMOS categories. The scales derived and the values assigned individual codes under each scale are delineated in the Definition of Scales contained in Appendix C; these are briefly summarized below. The selection of the instructional dimensions to be assessed was based on both a theory of independent component process in reading (see Volume 5), and knowledge about the characteristics of effective instruction. Given (1) the theory that reading consists of a set of relatively independent component processes, (2) a measurement system which has been designed to assess growth in each of these components (namely, the Interactive Reading Assessment System), and (3) a theory of what general characteristics of instruction are important in its effective delivery, then assessing the dimensions of instruction which theoretically should advance skill in these components is both natural and critical.

This derivation and assignment of raw RAMOS codes to ordinal scale values were based on the judgement of SEDL staff expert in both reading acquisition and elementary school instruction. Clearly, the theoretical base of the assignments may be open to question, but part of their assessment must be left to their usefulness in predicting reading growth (the subject of Volume 7). The instructional dimensions derived are summarized below:

Number of Students: the number of students contained in the instructional group.

Classification: the level of the instructor's formal training, ranging from minimal (volunteer) to mid-level (teacher aide) to substantial (substitute teacher, resource teacher, teacher).

Role: the level of formal instruction provided, ranging from minimal (preparation, control, management) to mid-level (facilitation) to substantial (direct instruction).

Subject Matter: the amount of reading generally required by the subject being taught, ranging from minimal (class business, art) to mid-level (science, mathematics) to substantial (reading).

Instructional Focus: the relative explicitness of the instructional emphases and strategies employed in three instructional subcategories:

Letter-Sound Unit: the relative explicitness of the instructional emphasis placed on decoding, ranging from work on isolated units (auditory discrimination, letter recognition, letter-name work) to non-explicit letter-sound pairing (whole word recognition, spelling practice) to explicit letter-sound pairing (letter cluster-sound recognition, letter-sound recognition, spelling pattern recognition).

Word Unit - Meaning: the relative explicitness of the instructional emphasis placed on word meaning, ranging from low (dictionary usage) to mid-level (noun derivative, compound words) to high (antonyms/synonyms, vocabulary enrichment).

Sentence and Text Units - Meaning: the relative explicitness of the instructional emphasis placed on sentence and text meaning, ranging from low (literal facts) to mid-level (story sequence, predicting events) to high (major ideas, making inferences).

Technique: the type of technique in which skills of visual or auditory pattern recognition are presented, as either parts-to-whole or whole-to-parts.

Language of Instruction: the language used in instruction delivery, ranging from all Spanish to alternating usage of English and Spanish to all English.

Materials (Primary and Ancillary): the amount of text contained in the materials used, ranging from minimal (art material, tape recorder) to mid-level (phrase card, chalkboard) to substantial (basal reader, library book).

Activity/Task: the level of formal language demand required by particular activity/tasks in three instructional subcategories:

Non-instructional: the type of instructional activity/task, as either non-instructional (clean-up, wait time) or instructional (all other activity/tasks). [Note: low values correspond to fewer instructional activity/tasks, high values to more instructional activity/tasks.]

Independent: the level of formal language demand for activity/tasks classified as independent work, ranging from minimal (art activity, copying material) to mid-level (writing from dictation, writing answers) to substantial (test taking, creating writing).

Listening and Responding in Group: the level of formal language demand for activity/tasks classified as listening and responding in groups, ranging from minimal (music activity, playing games) to mid-level (watch-listen, listen-story) to substantial (listen-lecture, discussion-speak).

Attention (Collection Years 1-2): the attention of the instructional group as rated relative to the activity/task required, ranging from low to medium to high.

Number of Nonengaged Students (Collection Years 3-5): the number of students contained in the instructional group who were not engaged in the activity/task being conducted.

Productivity: the rated productivity of the instructional group, ranging from low to medium to high.

Noise: the level of noise as rated relative to the activity/task required, ranging from low to medium to high.

In deriving these scale measures, each protocol contained in the raw RAMOS code computer files were "rectified" by a computer program that generated a line of characters for every minute of the observation period for each instantiated group, with each such line reflecting the codes in effect during that minute under each of the RAMOS categories. Thus, the raw RAMOS protocols, which the observer completed whenever a change in a RAMOS category occurred for a given group, were expanded by the program to reflect the minute-by-minute record of the groups under observation. The program generated a set of six such minute-by-minute descriptions, one for each of the six possible groups under observation, thus providing a minute-by-minute account of the status of each RAMOS category for each of a maximum of six groups instantiated at any given time. Since group data were not the subject of this study (although they are available), but rather, the data concerning individual target student instruction, the analysis program was designed to give a similar minute-by-minute account of the instruction received by individual targets. The program identified a given target student at the beginning of the observation (through an assigned unique identification number), and then tracked that student throughout the observation, regardless of the group (or more frequently, groups) the student may have been assigned to over the course of the observation period.

The "expansion" of the sample protocol given in Figure 1 will be used to illustrate these procedures. Table 1 identifies the contents of the column information contained in the sample raw code minute-by-minute expansions; Figure 2 depicts the expansion for target student 1 (unique identification number 5099) from the example protocol; and Figure 3 gives the same expansion for group 1 of that protocol (the interested reader is referred to Appendix D for the expansions generated for target students 2, 3, and 4, and for groups 2, 3, and 4).

The first 16 columns of numbers in Figures 2 and 3 provide identification information for determining the target student, teacher, and date of observation. In columns 20-21, the minutes of the observation are listed, and the remaining columns contain the raw values for each of the RAMOS categories. As shown in Figure 2, all category entries are constant for the first 28 minutes of the observation for this student, except for group size, which changed at minutes 8 and

Table 1

Variable Positions in RAMOS Expanded Raw Coding Analyses

COLUMNS	VARIABLE
01-04	STUDENT ID (Student analyses only)
05-07	TEACHER ID
08	GROUP ID (group analyses only)
11-16	DATE OF OBSERVATION
20-21	MINUTES
23	STATUS
24	GROUP ID
25-26	NUMBER OF STUDENTS
27	INSTRUCTOR: ID
28	INSTRUCTOR: CLASSIFICATION
29	INSTRUCTOR: ROLE
30	SUBJECT
31-33	INSTRUCTIONAL FOCUS
34	TECHNIQUE
35	LANGUAGE OF INSTRUCTION
36-37	PRIMARY MATERIALS
38-39	ANCILLARY MATERIALS
40-42	ACTIVITY-TASK
43-44	NUMBER NON-ENGAGED
45	PRODUCTIVITY
46	NOISE
47	ATTENTION
50	MOMENTARY CONTROL
53-54	OBSERVATION LENGTH IN MINUTES
57-58	OBSERVED LENGTH IN MINUTES

570

5099202	012483	1	N1190XXLSWSXECBFFWCMOOMLX	51
5099202	012483	2	N1190XXLSWSXECBFFWCMOOMLX	
5099202	012483	3	N1190XXLSWSXECBFFWCMOOMLX	
5099202	012483	4	N1190XXLSWSXECBFFWCMOOMLX	
5099202	012483	5	N1190XXLSWSXECBFFWCMOOMLX	
5099202	012483	6	N1190XXLSWSXECBFFWCMOOMLX	
5099202	012483	7	N1190XXLSWSXECBFFWCMOOMLX	
5099202	012483	8	N1140XXLSWSXECBFFWCMOOMLX	
5099202	012483	9	N1140XXLSWSXECBFFWCMOOMLX	
5099202	012483	10	N1140XXLSWSXECBFFWCMOOMLX	
5099202	012483	11	N1140XXLSWSXECBFFWCMOOMLX	
5099202	012483	12	N1140XXLSWSXECBPFWCMOOMLX	
5099202	012483	13	N1110XXLSWSXECBFFWCMOOMLX	
5099202	012483	14	N1110XXLSWSXECBPFWCMOOMLX	
5099202	012483	15	N1110XXLSWSXECBFFWCMOOMLX	
5099202	012483	16	N1110XXLSWSXECBFPWCMOOMLX	
5099202	012483	17	N1110XXLSWSXECBFFWCMOOMLX	
5099202	012483	18	N1110XXLSWSXECBPPWCMOOMLX	
5099202	012483	19	N1110XXLSWSXECBFPWCMOOMLX	
5099202	012483	20	N1110XXLSWSXECBFPWCMOOMLX	
5099202	012483	21	N1110XXLSWSXECBFFWCMOOMLX	
5099202	012483	22	N1110XXLSWSXECBFFWCMOOMLX	
5099202	012483	23	N1110XXLSWSXECBFPWCMOOMLX	
5099202	012483	24	N1110XXLSWSXECBPPWCMOOMLX	
5099202	012483	25	N1110XXLSWSXECBPPWCMOOMLX	
5099202	012483	26	N1110XXLSWSXECBPPWCMOOMLX	
5099202	012483	27	N1110XXLSWSXECBPPWCMOOMLX	
5099202	012483	28	N1110XXLSWSXECBPPWCMOOMLX	
5099202	012483	29	N4041TNRGCNPECHXXRRROOMLX	
5099202	012483	30	N4041TNRGCNPECHXXRRROOMLX	
5099202	012483	31	N4041TNRGCNPECHXXRRAOOMLX	
5099202	012483	32	N4041TNRGCNPECHXXRRROOMLX	
5099202	012483	33	N4041TNRGCNPECHXXRRROOMLX	
5099202	012483	34	N4041TNRGCNPECHXXRRROOMLX	
5099202	012483	35	N4041TNRGCNPECHXXRRROOMLX	
5099202	012483	36	N4041TNRGCNPECHXXRRROOMLX	
5099202	012483	37	N4041TNRDCSPEQTXXQSAOOMLX	
5099202	012483	38	N4041TFRSORXEBRXXRRAOOMLX	
5099202	012483	39	N4041TFRSORXEBRXXRRAOOMLX	
5099202	012483	40	N4041TFRSORXEBRXXRRAOOMLX	
5099202	012483	41	N4041TNRCLFXEBRQTQSAOOMLX	
5099202	012483	42	N4041TNRCLFXEBRQTQSAOOMLX	
5099202	012483	43	N4041TNRCLFXEBRQTQSAOOMLX	
5099202	012483	44	N4041TNRCLFXEBRQTQSAOOMLX	
5099202	012483	45	N4041TNRCLFXEBRQTQSAOOMLX	
5099202	012483	46	N4041TNRCLFXEBRQTQSAOOMLX	
5099202	012483	47	N4041TNRCLFXEBRQTQSAOOMLX	
5099202	012483	48	N4040XXLSWSXECBFFWCMOOMLX	
5099202	012483	49	N[illegible]0XXLSWSXECBFFWCMOOMLX	
5099202	012483	50	N4040XXLSWSXECBFFWCMOOMLX	
5099202	012483	51	N4040XXLSWSXECBFFWCMOOMLX	

Figure 2. **RAMOS expanded raw coding for target student 5099 based on sample protocol.**

2021	012483	1	N1190XXLSWSXECBPFWCMOOMLX	51 51
2021	012483	2	N1190XXLSWSXECBPFWCMOOMLX	
2021	012483	3	N1190XXLSWSXECBPFWCMOOMLX	
2021	012483	4	N1190XXLSWSXECBPFWCMOOMLX	
2021	012483	5	N1190XXLSWSXECBPFWCMOOMLX	
2021	012483	6	N1190XXLSWSXECBPFWCMOOMLX	
2021	012483	7	N1190XXLSWSXECBPFWCMOOMLX	
2021	012483	8	N1140XXLSWSXECBPFWCMOOMLX	
2021	012483	9	N1140XXLSWSXECBPFWCMOOMLX	
2021	012483	10	N1140XXLSWSXECBPFWCMOOMLX	
2021	012483	11	N1140XXLSWSXECBPFWCMOOMLX	
2021	012483	12	N1140XXLSWSXECBPFWCMOOMLX	
2021	012483	13	N1110XXLSWSXECBPFWCMOOMLX	
2021	012483	14	N1110XXLSWSXECBPFWCMOOMLX	
2021	012483	15	N1110XXLSWSXECBPFWCMOOMLX	
2021	012483	16	N1110XXLSWSXECBPFWCMOOMLX	
2021	012483	17	N1110XXLSWSXECBPFWCMOOMLX	
2021	012483	18	N1110XXLSWSXECBPFWCMOOMLX	
2021	012483	19	N1110XXLSWSXECBPFWCMOOMLX	
2021	012483	20	N1110XXLSWSXECBPFWCMOOMLX	
2021	012483	21	N1110XXLSWSXECBPFWCMOOMLX	
2021	012483	22	N1110XXLSWSXECBPFWCMOOMLX	
2021	012483	23	N1110XXLSWSXECBPFWCMO OMLX	
2021	012483	24	N1110XXLSWSXECBPFWCMOOMLX	
2021	012483	25	N1110XXLSWSXECBPFWCMOOMLX	
2021	012483	26	N1110XXLSWSXECBPFWCMOOMLX	
2021	012483	27	N1110XXLSWSXECBPFWCMOOMLX	
2021	012483	28	N1110XXLSWSXECBPFWCMOOMLX	
2021	012483	29	N1070XXLSWSXECBPFWCMOOMLX	
2021	012483	30	N1070XXLSWSXECBPFWCMOOMLX	
2021	012483	31	N1070XXLSWSXECBPFWCMOOMLX	
2021	012483	32	N1070XXLSWSXECBPFWCMOOMLX	
2021	012483	33	N1070XXLSWSXECBPFWCMOOMLX	
2021	012483	34	N1070XXLSWSXECBPFWCMOOMLX	
2021	012483	35	N1070XXLSWSXECBPFWCMOOMLX	
2021	012483	36	N1070XXLSWSXECBPFWCMOOMLX	
2021	012483	37	N1070XXLSWSXECBPFWCMOOMLX	
2021	012483	38	N1070XXLSWSXECBPFWCMOOMLX	
2021	012483	39	N1070XXLSWSXECBPFWCMOOMLX	
2021	012483	40	N1070XXLSWSXECBPFWCMOOMLX	
2021	012483	41	N1070XXLSWSXECBPFWCMOOMLX	
2021	012483	42	N1100XXLSWSXECBPFWCMOOMLX	
2021	012483	43	N1100XXLSWSXECBPFWCMOOMLX	
2021	012483	44	N1100XXLSWSX .CBPFWCMOOMLX	
2021	012483	45	N1100XXLSWSXECE. FWCMOOMLX	
2021	012483	46	N1100XXLSWSXECBPFWCMOOMLX	
2021	012483	47	N1100XXLSWSXECBPFWCMOOMLX	
2021	012483	48	N1101TFLSWSXECBPFWCMOOMLX	
2021	012483	49	N1101TFLSWSXECBPFWCMOOMLX	
2021	012483	50	N1101TFLSWSXECBPFWCMOOMLX	
2021	012483	51	N1101TFLSWSXECBPFWCMOOMLX	

Figure 3. RAMOS raw coding for group 1 based on sample protocol.

13. At minute 29, this target student was assigned to group 4 for reading instruction directed by the teacher. This instruction lasted until minute 48, and reveals changes in the teacher's role, instructional focus, technique, materials, and activity/task during these 19 minutes. At minute 49, the student was returned to the language arts activities originally assigned, and continued that work until the observation ended three minutes ter.

Figure 3, showing the raw code expansion for group 1, reveals no changes in any of the RAMOS categories with the exception of group size (changes in Number of Students at minutes 8, 13, 29, and 42) and those categories associated with the instructor (not present with the group until the last four minutes of the observation). Recall that these group data are not of primary import for this study, but rather the data based on individual target students -- they are discussed here only to show the contrast between student and group level analyses.

The next step taken in the analysis of these protocols was to convert the raw RAMOS codes contained in these minute-by-minute listings to their numeric assignments under each defined scale. Again, the process will be illustrated by reference to the raw code expansions described above for target student 1 and group 1 in the sample protocol. Table 2 identifies the contents of the column information contained in the sample numeric code conversions; Figure 4 depicts the coded expansion for target student 1 corresponding to the raw coding given for this student in Figure 2; and Figure 5 gives the same expansion for group 1 of that protocol corresponding to the raw coding given in Figure 3 (the coded expansions for the remaining target students and groups appearing in the sample protocol are included in Appendix D).

As in the raw code expansion figures, the first 16 columns of numbers in Figure 4 provide identification information for the target student, teacher, and date of observation. In columns 20-21, the minutes of the observation are listed (starting from minute 1), and the remaining columns to the right contain the coded values for the RAMOS scaled categories. In these categories a value of zero has special significance -- it does not indicate a low value (except for the number of students and the number of nonengaged students), but rather that scaling was not applicable. For example, for the Inst uctor Classification codes appearing in columns 28, a value of '0' appears for the first 28 minutes and the last 4 minutes, as no instructor was with this target student during these times. However, for the 19 minutes an instructor was present with this target student (minutes 29 through 47), a non-zero value reflecting the scaling of the instructor's classification appears. During this time, the teacher was providing direct instruction to the group to which this target student was assigned, and the scale value of '7' is found during these minutes, reflecting the high level of formal training associated with the instructor (a teacher). A similar pattern holds for the category of Instructor Role coded in column 29.

Table 2

Variable Positions in RAMOS Expanded Scale Coding Analyses

COLUMNS	VARIABLE
01-04	STUDENT ID (Student analyses only)
05-07	TEACHER ID
08	GROUP ID (group analyses only)
11-16	DATE OF OBSERVATION
20-21	MINUTES
25-26	NUMBER OF STUDENTS
28	INSTRUCTOR: CLASSIFICATION
29	INSTRUCTOR: ROLE
30	SUBJECT
32	INSTRUCTIONAL FOCUS: LETTER-SOUND UNIT
33	INSTRUCTIONAL FOCUS: WORD UNIT
34	INSTRUCTIONAL FOCUS: SENTENCE/TEXT UNIT
36	TECHNIQUE
37	LANGUAGE OF INSTRUCTION
38	PRIMARY MATERIALS
39	ANCILLARY MATERIALS
41	ACTIVITY-TASK: NON-INSTRUCTIONAL
42	ACTIVITY-TASK: INDEPENDENT
43	ACTIVITY-TASK: LISTENING/RESPONDING IN GROUP
45-46	NUMBER NON-ENGAGED
47	PRODUCTIVITY
48	NOISE
49	ATTENTION
52	MOMENTARY CONTROL
55-56	OBSERVATION LENGTH IN MINUTES
59-60	OBSERVED LENGTH IN MINUTES

574

BEST COPY AVAILABLE

5099202	012483	1	19	008	001	0355	210	0320	0	51	51
5099202	012483	2	19	008	001	0355	210	0320	0		
5099202	012483	3	19	008	001	0355	210	0320	0		
5099202	012483	4	19	008	001	0355	210	0320	0		
5099202	012483	5	19	008	001	0355	210	0320	0		
5099202	012483	6	19	008	001	0355	210	0320	0		
5099202	012483	7	19	008	001	0355	210	0320	0		
5099202	012483	8	4	008	001	0355	210	0320	0		
5099202	012483	9	14		001	0355	210	0320	0		
5099202	012483	10	14	.	001	0355	210	0320	0		
5099202	012483	11	14	008	001	0355	210	0320	0		
5099202	012483	12	14	008	001	0355	210	0320	0		
5099202	012483	13	11	008	001	0355	210	0320	0		
5099202	012483	14	11	008	001	0355	210	0320	0		
5099202	012483	15	11	008	001	0355	210	0320	0		
5099202	012483	16	11	008	001	0355	210	0320	0		
5099202	012483	17	11	008	001	0355	210	0320	0		
5099202	012483	18	11	008	001	0355	210	0320	0		
5099202	012483	19	11	008	001	0355	210	0320	0		
5099202	012483	20	11	008	001	0355	210	0320	0		
5099202	012483	21	11	008	001	0355	210	0320	0		
5099202	012483	22	11	008	001	0355	210	0320	0		
5099202	012483	23	11	008	001	0355	210	0320	0		
5099202	012483	24	11	008	001	0355	210	0320	0		
5099202	012483	25	11	008	001	0355	210	0320	0		
5099202	012483	26	11	008	001	0355	210	0320	0		
5099202	012483	27	11	008	001	0355	210	0320	0		
5099202	012483	28	11	008	001	0355	210	0320	0		
5099202	012483	29	4	799	200	1340	203	0320	0		
5099202	012483	30	4	799	200	1340	203	0320	0		
5099202	012483	31	4	799	200	1340	202	0320	0		
5099202	012483	32	4	799	200	1340	203	0320	0		
5099202	012483	33	4	799	200	1340	203	0320	0		
5099202	012483	34	4	799	200	1340	203	0320	0		
5099202	012483	35	4	799	200	1340	203	0320	0		
5099202	012483	36	4	799	200	1340	203	0320	0		
5099202	012483	37	4	799	300	1330	203	0320	0		
5099202	012483	38	4	769	200	0380	202	0320	0		
5099202	012483	39	4	769	200	0380	202	0320	0		
5099202	012483	40	4	769	200	0380	202	0320	0		
5099202	012483	41	4	799	001	0383	203	0320	0		
5 99202	012483	42	4	799	001	0383	203	0320	0		
5099202	012483	43	4	799	001	0383	203	0320	0		
5099202	012483	44	4	799	001	0383	203	0320	0		
5099202	012483	45	4	799	001	0383	203	0320	0		
5099202	012483	46	4	799	001	0383	203	0320	0		
5099202	012483	47	4	799	001	0383	203	0320	0		
5099202	012483	48	4	008	001	0355	210	0320	0		
5099202	012483	49	4	008	001	0355	210	0320	0		
5099202	012483	50	4	008	001	0355	210	0320	0		
5099202	012483	51	4	008	001	0355	210	0320	0		

Figure 4. RAMOS expanded scale coding for target student 5099 based on sample protocol.

2021	012483	1	19	008	001	0355	210	0320	1	31	51
2021	012483	2	19	008	001	0355	210	0320	0		
2021	012483	3	19	008	001	0355	210	0320	0		
2021	012483	4	19	008	001	0355	210	0320	0		
2021	012483	5	19	008	001	0355	210	0320	0		
2021	012483	6	19	008	001	0355	210	0320	0		
2021	012483	7	19	008	001	0355	210	0320	0		
2021	012483	8	14	008	001	0355	210	0320	0		
2021	012483	9	14	008	001	0355	210	0320	0		
2021	012483	10	14	008	001	0355	210	0320	0		
2021	012483	11	14	008	001	0355	210	0320	0		
2021	012483	12	14	008	001	0355	210	0320	0		
2021	012483	13	11	008	001	0355	210	0320	0		
2021	012483	14	11	008	001	0355	210	0320	0		
2021	012483	15	11	008	001	0355	210	0320	0		
2021	012483	16	11	008	001	0355	210	0320	0		
2021	012483	17	11	008	001	0355	210	0320	0		
2021	012483	18	11	008	001	0355	210	320	0		
2021	012483	19	11	008	001	0355	210	0320	0		
2021	012483	20	11	008	001	0355	210	0320	0		
2021	012483	21	11	008	001	0355	210	0320	0		
2021	012483	22	11	008	001	0355	210	0320	0		
2021	012483	23	11	008	001	0355	210	0320	0		
2021	012483	24	11	008	001	0355	210	0320	0		
2021	012483	25	11	008	001	0355	210	0320	0		
2021	012483	26	11	008	001	0355	210	0320	0		
2021	012483	27	11	008	001	0355	210	0320	0		
2021	012483	28	11	008	001	0355	210	0320	0		
2021	012483	29	7	008	001	0355	210	0320	0		
2021	012483	30	7	008	001	0355	210	0320	0		
2021	012483	31	7	008	001	0355	210	0320	0		
2021	012483	32	7	008	001	0355	210	0320	0		
2021	012483	33	7	008	001	0355	210	0320	0		
2021	012483	34	7	008	001	0355	210	0320	0		
2021	012483	35	7	008	001	0355	210	0320	0		
2021	012483	36	7	008	001	0355	210	0320	0		
2021	012483	37	7	008	001	0355	210	0320	0		
2021	012483	38	7	008	001	0355	210	0320	0		
2021	012483	39	7	008	001	0355	210	0320	0		
2021	012483	40	7	008	001	0355	210	0320	0		
2021	012483	41	7	008	001	0355	210	0320	0		
2021	012483	42	10	008	001	0355	210	0320	0		
2021	012483	43	10	008	001	0355	210	0320	0		
2021	012483	44	10	008	001	0355	210	0320	0		
2021	012483	45	10	008	001	0355	210	0320	0		
2021	012483	46	10	008	001	0355	210	0320	0		
2021	012483	47	10	008	001	0355	210	0320	0		
2021	012483	48	10	768	001	0355	210	0320	0		
2021	012483	49	10	768	001	0355	210	0320	0		
2021	012483	50	10	768	001	0355	210	0320	0		
2021	012483	51	10	768	001	0355	210	0320	0		

Figure 5. RAMOS expanded scale coding for group 1 based on sample protocol.

Continuing with the example given in Figure 4, the three subscales derived from the raw codes under the category of Instructional Focus (namely, Letter-Sound Unit, Word Unit - Meaning, and Sentence and Text Unit - Meaning) appear in columns 32-34, respectively. For instruction pertaining to decoding (Letter-Sound Unit), the only applicable codes appear during the first 12 minutes the teacher is instructing the group (minutes 29 through 40): the first 8 minutes of 'GCN' receive mid-scale values of '2' (work on contractions represents relatively non-explicit instruction in letter-sound pairing); the following one minute of 'DCS' receives a top scale value of '3' (work on letter cluster-sound recognition represents explicit instruction in letter-sound pairing); and the final three minutes of 'SOR' receive mid-scale values of '2' (work on oral reading fluency represents instruction which is relatively non-explicit in letter-sound pairing).

In column 33 of Figure 4, the category of instruction devoted to word meanings (Word Unit - Meaning) is coded, and it is always represented as zero in this sample observation. This indicates that of all the instructional focus codes appearing in the expanded raw coding of this observation for this student, no code was found that applied to this subscale of instruction -- during this observation, this student received no instruction related to the teaching of the meanings of words.

The reader is left to track the coding of the other instructional dimensions for target student 1, and also for group 1 (identically structured) as displayed in Figure 5. What is important to keep in mind at this point is that the expanded, scaled RAMOS protocols, as displayed in Figure 4, represent a set of instructional dimensions with respect to both their quality (reflected in the relative magnitude of the scaled values) and quantity (reflected in the number of such scaled values). All that remains is to reduce such expanded protocols to a smaller, more manageable, set of measures which can retain these features of quality and quantity.

The final step involved in summarizing a given student's instruction during a given protocol was to compute a set of indices based on the expanded scaled codings just described. First, the mean value of all applicable (i.e., non-zero) scale values within a given scale was computed as an index of the quality of that instructional dimension. Second, the standard deviation of the scaled values around the mean value was computed as an index of the variability of the quality of instruction under the instructional dimension. Finally, the percent of time in which applicable codes were found, relative to the total amount of time the student was observed, was computed as an index of the quantity of the instructional dimension found during the observation.

In illustrating this procedure, these sets of summary indices for the four target students identified in the sample protocol are displayed in Figure 6; Figure 7 displays similar information based on the our groups realized during the observation.

STD ID	TCH ID	RCRD NUM	DATE	NUM STDS	INST CLSS	INST ROLE	SUBJ	INST FOCUS: LET-SND	WORD	SENT TEXT	TECH	LANG	MAT: PRIM	MAT: AMCL	ACTIVITY-TASK: NON-INST	INDP	LSTN RESP	NON-ENGD	PRDC	NOISE	ATTN	MMNT CNTR	TOTL OBSR	STD OBSR
5099	202	01	012483	9.2	7.0	8.5	8.4	2.1	-1.0	1.0	1.0	3.0	5.4	4.6	2.0	1.0	2.8	0.0	3.0	2.0	-1.0	-1.0	51	51
5099	202	02	012483	3.4	0.0	1.1	0.5	0.3	-1.0	0.0	0.0	0.0	1.4	0.8	0.0	0.0	0.4	0.0	0.0	0.0	-1.0			
5099	202	03	012483	99.9	37.3	37.3	99.9	23.5	-1.0	76.5	17.6	99.9	99.9	76.5	99.9	62.7	37.3	99.9	99.9	99.9	-1.0			
5109	202	01	012483	6.9	7.0	7.1	8.4	2.1	-1.0	1.0	1.0	3.0	5.9	4.8	2.0	1.0	2.3	0.0	3.0	2.0	-1.0	1.0	51	51
5109	202	02	012483	4.9	0.0	1.4	0.5	0.4	-1.0	0.0	0.0	0.0	1.5	0.7	0.0	0.0	0.5	0.0	0.3	0.0	-1.0			
5109	202	03	012483	99.9	41.2	41.2	99.9	27.5	-1.0	66.7	9.8	99.9	94.1	56.7	99.9	58.8	41.2	99.9	96.1	99.9	-1.0			
5120	202	01	012483	6.9	7.0	7.1	8.4	2.1	-1.0	1.0	1.0	3.0	5.9	4.8	2.0	1.0	2.3	0.0	3.0	2.0	-1.0	1.0	51	51
5120	202	02	012483	4.9	0.0	1.4	0.5	0.4	-1.0	0.0	0.0	0.0	1.5	0.7	0.0	0.0	0.5	0.0	0.3	0.0	-1.0			
5120	202	03	012483	99.9	41.2	41.2	99.9	27.5	-1.0	66.7	9.8	99.9	94.1	66.7	99.9	58.8	41.2	99.9	96.1	99.9	-1.0			
5259	202	01	012483	3.7	7.0	9.0	8.0	2.0	2.0	1.0	-1.0	3.0	4.9	4.7	2.0	1.0	3.0	0.0	3.0	2.0	-1.0	-1.0	51	51
5259	202	02	012483	0.7	0.0	0.0	0.0	0.0	0.0	0.0	-1.0	0.0	0.3	0.7	0.0	0.0	0.0	0.0	0.0	0.0	-1.0			
5259	202	03	012483	99.9	13.7	13.7	99.9	9.8	3.9	86.3	-1.0	99.9	99.9	99.9	99.9	86.3	13.7	99.9	99.9	99.9	-1.0			

Figure 6. RAMOS derived summary indices for four target students from the Site 5 second grade protocol.

								INST FOCUS:							ACTIVITY-TASK:									
TCH ID	GRP ID	RCRD NUM	DATE	NUM STDS	INST CLSS	INST ROLE	SUBJ	LET-SND	WORD	SENT TEXT	TECH	LANG	MAT: PRIM	MAT: ANCL	NON-INST	INDP	LSTN RESP	NON-ENGD	PRDC	NOISE	ATTN	MNNT CNTR	TOTL OBSR	STD OBSR
202	1	01	012483	11.2	7.0	6.0	8.0	-1.0	-1.0	1.0	-1.0	3.0	5.0	5.0	2.0	1.0	-1.0	0.0	3.0	2.0	-1.0	-1.0	51	51
202	1	02	012483	3.8	0.0	0.0	0.0	-1.0	-1.0	0.0	-1.0	0.0	0.0	0.0	0.0	0.0	-1.0	0.0	0.0	0.0	-1.0			
202	1	03	012483	99.9	7.8	7.8	99.9	-1.0	-1.0	99.9	-1.0	99.9	99.9	99.9	99.9	99.9	-1.0	99.9	99.9	99.9	-1.0			
202	2	01	012483	5.7	7.0	9.0	8.0	2.0	2.0	1.0	-1.0	3.0	4.9	4.7	2.0	1.0	3.0	0.0	3.0	2.0	-1.0	-1.0	51	51
202	2	02	012483	0.7	0.0	0.0	0.0	0.0	0.0	0.0	-1.0	0.0	0.3	0.7	0.0	0.0	0.0	0.0	0.0	0.0	-1.0			
202	2	03	012483	99.9	13.7	13.7	99.9	9.8	3.9	86.3	-1.0	99.9	99.9	99.9	99.9	86.3	13.7	99.9	99.9	99.9	-1.0			
202	3	01	012483	5.0	7.0	7.1	8.5	2.1	-1.0	1.0	1.0	3.0	6.1	4.7	2.0	1.0	2.3	0.0	3.0	2.0	-1.0	1.0	51	44
202	3	02	012483	0.0	0.0	1.4	0.5	0.4	-1.0	0.0	0.0	0.0	1.6	0.7	0.0	0.0	0.5	0.0	0.3	0.0	-1.0			
202	3	03	012463	99.9	47.7	47.7	99.9	31.8	-1.0	61.4	11.4	99.9	93.2	61.4	99.9	52.3	47.7	99.9	95.5	99.9	-1.0			
202	4	01	012483	4.0	7.0	8.5	8.8	2.1	-1.0	1.0	1.0	3.0	5.9	3.7	2.0	1.0	2.8	0.0	3.0	2.0	-1.0	-1.0	51	23
202	4	02	012483	0.0	0.0	1.1	0.4	0.3	-1.0	0.0	0.0	0.0	2.0	1.0	0.0	0.0	0.4	0.0	0.0	0.0	-1.0			
202	4	03	012483	99.9	82.6	82.6	99.9	52.2	-1.0	47.8	39.1	99.9	99.9	47.8	99.9	17.4	82.6	99.9	99.9	99.9	-1.0			

Figure 7. RAMOS derived summary indices for four groups from the Site 5 second grade sample protocol.

530 531

Treating Figure 6 first, the first four columns again give student-protocol identification information (student identification number under STD ID, teacher identification number under TCH ID, record number under RCRD NUM, and observation date under DATE). The remaining columns to the right cover the RAMOS scales, and are presented under mnemonic headings.

For the first target student, Figure 7 displays values of 11.2, 3.8, and 99.9 under the category of Number of Students (NUM STDS), representing the mean (record number 1), standard deviation (record number 2), and percent of time (record number 3) which summarize the size of the group this student was associated with during the observation. Thus, the average group size was about 11 students; the standard deviation of 3.8 students indicates that the group size had a fair degree of fluctuation, and the percent value of 99.9 indicates that group size information was available for this student for the entire period observed (due to space limitations, values of 100% were entered in the data files as '99.9', but such values were converted back to 100% in all analysis programs).

Recalling that the instructor was only associated with target student 1 during a 19 minute period, the summary indices displayed in Figure 6 for this student under Instructor Classification (INST CLSS) are the following: (1) mean value of 7.0 (indicating that whenever an instructor was present with this student during this observation, that instructor's average level of formal training was high), (2) standard deviation of 0.0 (revealing that the level of formal training of any instructors associated with this student during this observation did not vary), and (3) time percentage of 37.3 (showing that this student was with some instructor for slightly more than a third of the observation -- 19 minutes observed with an instructor divided by 51 minutes of observed time).

As a third example, recall that no applicable codes relating to instruction dealing with word meanings were encountered in the sample protocol for target student 1 -- the relevant entries for this student in Figure 6 (student 5099 under INST FOCUS: WORD) are all '-1.0', demonstrating the use of this value to mark cases where no applicable codes for a given scale were found during the time the student was observed. A similar use of this coding is seen under the category for Attention (ATTN); here all students and groups receive values of '-1.0' since the sample protocol comes from the Year 5 data collection period which did not use this category, instead employing the more specific Number of Nonengaged Students.

Two additional categories are important to note. First, the number of momentary controls observed (discussed earlier) is represented as a simple sum. As shown in Figure 6 under MMNT CNTR, appropriate values of '1.0' are entered for the two students (5109 and 5120) assigned to the group affected by the single momentary control found in the sample protocol (entries of '-1.0' for the other two students signify that no such controls were found). Second, the total length of the observation in minutes, as well as the number of minutes the target student (or group, depending upon the analysis) was observed, is calculated. Figure 6 shows both these values to be 51 minutes for the four target students. However, the group data presented in Figure 7, show that only the first two groups existed throughout the observation, with groups 3 and 4, formed for only 44 and 23 minutes, respectively.

Again, the interested reader is left to explore the correspondence between the remaining summary indices and the relevant expanded scale codings. Such summary indices were computed for each of the RAMOS scales described above, and provide estimates of both the quality and quantity of the defined dimensions of instruction. The use of the term "quality" here does not imply any evaluation of the appropriateness of the instruction -- the skills of the students in a given group may be such that certain types of instruction are obviated. However, this information provides a basis for assessing the kind of instruction received (i.e., its quality and quantity relative to the dimensions defined in this study), and subsequent analyses will provide assessments of whether or not instruction so defined influences relevant reading skills of these students.

Finally, it should be noted that while both the raw and coded expansions described above were available from the analysis program created to compute the summary indices, the actual computing algorithm did not require such expansions be made -- they primarily serve a pedagogical function in explicating the meaning of the derived summary indices.

The summary indices just described are the basic measures of instruction employed in this study, yet they are not the end-product used in the analyses integrating reading skill and instruction. The derivation of these measures is discussed next.

Aggregation Procedures

The average scale values obtained from individual protocols (as just described) were next themselves averaged <u>within semester</u> for each individual target student (i.e., all protocol averages for a given student within a given semester were collected, then averaged). Finally, the Fall and Spring average semester values were averaged to obtain a yearly average value for each target student for each scale (thus giving equal weighting to Fall and Spring observations in cases where the numbers of observations, or lengths of observation for a given student, were not equally distributed across the two semesters). The interpretations of the resulting quality and quantity values are not changed: for a given scale, the quality index represents the average scale value and the quantity index represents the average percent of student time spent in the instructional dimension relevant to that scale (with both indices based on equal weightings of Fall and Spring observations). These values are the basis of most of the subsequent analyses.

In computing these averages over individual observation summaries, it is important to note that all instructional focus and activity/task percentage values of '-1.0' (indicating that the observation contained no codes applicable to the particular scale) were converted to 0.0% for the purpose of computing an average over the set of observations for a given semester. For instance, if codes relevant to the IFWD scale were not encountered in one observation, but were encountered in 20% of a second observation, then averaging these percentages across the two observations would result in a value of 10% rather than the 20% value that would result if the value of '-1.0' in the first observation was considered missing.

One final procedural point needs to be noted. In general, students were assigned to a single teacher for reading instruction throughout the year. For the majority of cases where this was not true, changes in reading teachers occurred at semester break upon return from the Christmas holidays. For the few remaining cases, changes occurred at the beginning of the year (usually in kindergarten), as class assignments in general were settled; observations for these initial reading teachers were dropped from subsequent analyses. Thus, in general, all semester averages for a given target student were based on observations for a single teacher -- only in the rare cases where teachers left the school during the year and were subsequently replaced, were observations from different teachers averaged within semester.

Reliability

The interrater reliability of the RAMOS was assessed through informal means. At the end of the first intensive training session conducted during the first year of a site's entry into the study, the data collectors together observed a number of classroom reading periods. Their protocols were compared at the end of the day, and in general, few major discrepancies were found; those that did appear were discussed and a standard handling of any divergent coding was agreed upon. As mentioned above, data collectors maintained close contact during the year, both among themselves and with SEDL staff, thus facilitating communication about any problematic cases encountered.

Descriptive Statistics

In this section, the descriptive statistics on the RAMOS summary indices for the bilingual sample are described, treating English first, then Spanish.

Instruction in English Reading

In presenting the English reading observation summaries, first the aggregate data are discussed for the entire bilingual sample, then the differences in instruction found for individual sites.

Aggregate Descriptive Data

For the English reading protocols, Tables 3 and 4 display the descriptive statistics (mean, standard deviation, and number of cases) for the average quality and quantity indices of instruction, respectively. The data are summarized for each of the five instructional years, and the instructional scales appear under mnemonic headings. For the quality indices (Table 3), the following four-character scale names are used (the characters MN appended to these names stand for Mean):

- NSTD: Number of Students
- CLSS: Instructor: Classification
- ROLE: Instructor: Role
- SBJC: Subject
- IFLT: Instructional Focus: Letter-Sound Unit
- IFWD: Instructional Focus: Word Unit - Meaning
- IFST: Instructional Focus: Sentence/Text Unit - Meaning
- TECH: Technique
- LANG: Language of Instruction
- MAT1: Primary Materials
- MAT2: Ancillary Materials
- ATNT: Activity/Task: Non-instructional
- ATIN: Activity/Task: Independent
- ATLR: Activity/Task: Listening and Responding in Group
- NEN2: Number of Nonengaged Students (Version 2, Years 3-5)
- PRDC: Productivity
- NOIS: Noise
- ATT1: Attention (Version 1, Years 1-2)
- CTRL: Number of Momentary Controls
- OBTM: Observation Time
- STTM: Student Observed Time

For the quantity indices (Table 4), the same four-character scale names are used (the characters PR appended to these names stand for Percentage), but the following additional scale values are included. First, IFTTPR is the total average percentage of time in which an applicable instructional focus code was found (i.e., it is the averaged total of the percent of time associated with the three scales derived from the instructional focus codes: IFLTPR, IFWDPR, and IFSTPR). As such, it represents the percent of time during which the instruction observed is focused on literacy skills. Second, IFT2PR is a similar average, but leaving out IFLTPR -- it represents the average percent of literacy-focused instructional time which is not devoted to decoding instruction. Similarly, ATTTPR represents the total average percentage of time in which an instructional activity/task was conducted -- it is the averaged total of the percentage of time associated with the two scales derived from the activity/task codes, ATINPR and ATLRPR.

It is important to note that the instructional focus and activity/task total percentages (namely, IFTTPR and ATTTPR) will not necessarily sum to 100%. When they do, it indicates that for the

Table 3

Reading and Mathematics Observation System - English:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for the Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
NSTDMN	M	12.1	13.5	15.3	15.2	14.9
NSTDMN	S	4.9	4.7	4.3	3.8	4.9
NSTDMN	N	140	215	224	94	59
CLSSMN	M	6.5	6.5	6.7	6.7	6.8
CLSSMN	S	0.5	0.6	0.4	0.4	0.4
CLSSMN	N	140	212	222	94	59
ROLEMN	M	6.7	6.1	6.3	6.0	5.9
ROLEMN	S	1.3	1.3	1.5	1.3	1.4
ROLEMN	N	140	212	222	94	59
SBJCMN	M	6.8	7.9	8.2	8.[illegible]	7.9
SBJCMN	S	1.7	0.7	0.8	0.7	0.9
SBJCMN	N	140	215	224	94	59
IFLTMN	M	1.6	2.0	2.0	2.0	2.0
IFLTMN	S	0.5	0.3	0.2	0.2	0.1
IFLTMN	N	138	210	220	94	58
IFWDMN	M	2.8	2.8	2.5	2.4	2.8
IFWDMN	S	0.6	0.3	0.5	0.7	0.5
IFWDMN	N	71	90	144	92	54
IFSTMN	M	2.1	1.4	1.5	1.8	1.6
IFSTMN	S	0.5	0.3	0.4	0.4	0.5
IFSTMN	N	84	200	211	93	58
TECHMN	M	1.2	1.2	1.3	1.7	1.7
TECHMN	S	0.3	0.3	0.4	0.2	0.3
TECHMN	N	124	177	204	94	59
LANGMN	M	2.8	2.9	3.0	2.9	3.0
LANGMN	S	0.3	0.2	0.1	0.2	0.1
LANGMN	N	140	215	224	94	59
MAT1MN	M	4.0	5.6	6.0	6.7	6.6
MAT1MN	S	[illegible]	0.9	0.8		1.0
MAT1MN	N	140	215	224		59
MAT2MN	M	3.3	4.2	4.5		4.5
MAT2MN	S	1.4	0.9	0.9		0.8
MAT2MN	N	119	210	220	94	59
ATNTMN	M	1.9	1.9	1.9	1.9	1.9
ATNTMN	S	0.1	0.1	0.1	0.1	0.1
ATNTMN	N	140	215	224	94	59
ATINMN	M	1.6	1.5	1.7	1.9	2.0
ATINMN	S	0.6	0.4	0.4	0.3	0.3
ATINMN	N	102	212	223	94	56
ATLRMN	M	2.0	2.2	2.4	2.3	2.3
ATLRMN	S	0.5	0.4	0.4	0.3	0.5
ATLRMN	N	138	210	219	93	59

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
NEXZMN	M	1.4	1.7	1.5	2.1	2.7
NEXZMN	S	1.0	1.4	1.5	1.2	1.5
NEXZMN	N	91	159	204	94	59
PROCMN	M	3.4	3.0	3.0	3.3	3.1
PROCMN	S	0.4	0.3	0.4	0.4	0.5
PROCMN	N	135	214	223	94	59
NOISMN	M	2.5	2.4	2.3	2.2	2.4
NOISMN	S	0.4	0.4	0.3	0.4	0.5
NOISMN	N	140	215	224	94	59
ATT1MN	M	3.4	3.4	3.7		
ATT1MN	S	0.4	0.3	0.2		
ATT1MN	N	49	59	23		
CTRLMN	M	0.3	0.4	0.4	0.3	0.2
CTRLMN	S	0.5	0.8	0.5	0.4	0.3
CTRLMN	N	140	215	224	94	59
OBTMMN	M	49.3	45.2	46.3	42.6	40.4
OBTMMN	S	10.9	9.7	11.3	9.3	6.7
OBTMMN	N	140	215	224	94	59
STTMMN	M	46.7	41.2	43.0	41.6	39.3
STTMMN	S	11.1	8.3	11.4	9.3	7.2
STTMMN	N	140	215	224	94	59

586

Table 4

Reading and Mathematics Observation System - English: Descriptive Statistics on the Quantity Indices of Instruction by Instructional Year for the Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
CLSSPR	M	79.0	66.0	67.4	65.2	67.1
CLSSPR	S	19.6	23.0	23.2	20.8	21.0
CLSSPR	N	140	215	224	94	59
ROLEPR	M	79.2	66.4	70.3	66.6	70.7
ROLEPR	S	19.2	21.0	20.6	19.5	19.7
ROLEPR	N	140	212	222	94	59
IFLTPR	M	49.4	52.3	57.3	44.6	29.5
IFLTPR	S	25.5	20.3	18.8	17.3	18.0
IFLTPR	N	140	215	224	94	59
IFWDPR	M	9.1	4.8	6.9	14.3	20.1
IFWDPR	S	12.8	8.9	9.7	12.1	12.4
IFWDPR	N	140	215	224	94	59
IFSTPR	M	9.3	28.7	27.3	30.0	36.5
IFSTPR	S	10.0	21.2	18.5	16.7	18.3
IFSTPR	N	140	215	224	94	59
MAT1PR	M	88.0	93.6	93.9	93.2	91.1
MAT1PR	S	10.4	6.7	7.5	5.8	7.8
MAT1PR	N	140	215	224	94	59
MAT2PR	M	48.0	65.2	62.3	72.1	74.5
MAT2PR	S	26.9	17.3	22.1	15.8	20.0
MAT2PR	N	119	210	220	94	59
ATINPR	M	21.0	49.6	46.8	37.6	33.3
ATINPR	S	19.2	24.7	19.8	16.1	19.7
ATINPR	N	140	215	224	94	59
ATLRPR	M	67.1	41.9	45.6	51.7	53.0
ATLRPR	S	23.2	22.7	18.6	18.9	21.1
ATLRPR	N	140	215	224	94	59
IFTTPR	M	67.8	85.8	91.4	88.9	86.1
IFTTPR	S	29.3	14.1	9.4	7.7	10.4
IFTTPR	N	140	215	224	94	59
IFT2PR	M	18.4	33.5	34.1	44.3	56.6
IFT2PR	S	18.6	19.9	19.5	17.3	20.5
IFT2PR	N	140	215	224	94	59
ATTTPR	M	88.1	91.5	92.4	89.3	86.3
ATTTPR	S	13.9	7.2	8.8	6.6	9.7
ATTTPR	N	140	215	224	94	59
IFLTRP	M	74.6	61.2	63.0	50.1	35.1
IFLTRP	S	24.3	21.9	20.4	19.4	22.5
IFLTRP	N	139	215	224	94	59
IFWDRP	M	43.1	18.2	22.1	32.3	35.2
IFWDRP	S	31.0	26.6	26.3	23.2	19.4
IFWDRP	N	90	203	218	94	58
ATLRRP	M	75.5	46.4	49.6	57.5	60.9
ATLRRP	S	21.5	25.2	20.6	19.1	23.1
ATLRRP	N	138	215	224	94	59
STTMRT	M	95.3	93.0	93.4	97.9	97.4
STTMRT	S	10.7	14.0	12.5	5.7	8.0
STTMRT	N	140	215	[illegible]	94	59

597

entire observation, applicable codes were found under the relevant sets of scales. Conversely, the degree to which the sum differs from 100% is the degree to which applicable codes were not encountered, these generally representing cases where no instructional focus or activity/task was evident (coded 'X' by the observer).

The last set of quantity variables appearing in Table 4 represents relative percentages, percentage values based on the ratio of two percentage values (these variable labels are appended with RP). First, IFLTRP is the percentage of decoding time (IFLTPR) relative to the percentage of literacy-focused instructional time (IFTTPR). Second, IFWDRP is the percentage of time devoted to word meanings (IFWDPR) relative to the percentage of non-decoding time (IFT2PR, the percentage of time spent on word/sentence/text meaning). Third, ATLRRP is the percentage of time devoted to group instruction (ATLRPR) relative to the total percentage of time devoted to instructional activities (ATTTPR). Finally, STTMRT is the percentage of time the student was observed during an observation relative to the total observation time.

In describing the English reading classroom instruction as represented in these tables, only instructional years 1 through 4 will be considered since these data are the main focus of subsequent analyses. Note that the number of observations at each year reflects the cohort structure of the study (see Volume 2), and thus, the number of data points at instructional years 3 and 4 is reduced from that available at years 1 and 2, requiring caution in their interpretation. Further, the trends discussed over the instructional years have not been subjected to the usual statistical analyses -- these descriptions are intended only to give the reader a sense of the dimensions of the reading instruction provided to the target bilingual sample.

First, considering the number of students associated with the observed instructional groups of the target students, the trend appears to be for slightly smaller groups in the early years (around 13 students), and somewhat larger groups in later years (around 15 students).

For the instructor variables, for about two-thirds of the instruction observed, some instructor was associated with the target students. This instructor generally was a teacher, and the role played tended to be one of facilitation. The subject being taught generally required substantial reading, but the value suggests that both reading and language arts instruction took place simultaneously during the observation (i.e., one group in reading instruction, a second group in language arts instruction).

For the instructional focus variables, about 90% of the observed instruction contained instructional foci. About half the time was devoted to decoding instruction during the first and second instructional years, falling to about 30% by the fourth year. This instruction tended to be focused on isolated units in kindergarten, then on non-explicit letter-sound pairings for the remaining grades -- the relatively low variability in this measure suggests that little explicit letter-sound work was observed. The amount of time spent on word meanings was quite small during the first two years (around 5%), increasing to about 20% by the fourth year. The instruction that was offered, however, tended to be quite explicit. About 30% of the instructional focus time was on instruction in the meanings of sentences and texts for the first three years, with a slight increase in the fourth year. The quality of this instruction was fairly stable across the years, and was generally non-explicit (e.g., favoring a focus on literal facts over making inferences).

For materials, primary materials were used about 90% of the time, showing little change over the four years. The types of materials used, however, did change over time, showing an increase toward more text material. Ancillary materials were used about 60% of the time during the first two instructional years, and about 70% for years 3 and 4. The quality index of these materials was stable over the four instructional years, and was lower than the primary materials (e.g., tending more toward paper and pencil than text).

For the activity/task scales, about 90% of the observed time contained instructional activity/tasks. Independent work accounted for about half the instructional time during the first two years, dropping to about 35% in the following two years. Conversely, group work represented about 40% of the instructional time during the first two years, increasing to about 50% for the last two years. For independent work, the level of formal language demand increased over the instructional years, starting at a relatively low level in the first year, and increasing to mid-level by the last two years. For group work, however, little change in level of formal language demand occurred; it was, however, noticeably higher than that associated with independent work.

For the student response measures, the number of nonengaged students was low, with a slight increase in the last year; productivity was rated as medium in each year, while noise tended to be low. Finally, the number of momentary controls was generally low, with a slight drop over the four instructional years.

Site Differences

Given this overall description of the English reading observation data, individual site differences are discussed next. The quality and quantity summary measures for each site appear in Appendix E, tabled in the same format represented in Tables 3 and 4. The site differences discussed below have not been subjected to statistical test -- they are given only as supplemental information to the site description narratives provided in Volume 2.

For the number of students associated with the observed instructional groups of the target students, each of the five sites showed the general trend for slightly smaller groups in the early years and somewhat larger groups in later years. Site 0 showed the smallest average group sizes in each instructional year (ranging from about 10 to 14 students). The largest group sizes were generally found at Site 2, with the average group size found at first grade (about 21 students) being the largest value found at any site for any year.

Concerning the instructor variables, Site 1 showed the lowest percentage of instructional time with an associated instructor -- about 55% over the four instructional years (which is about 10% lower than the overall average). The associated instructor's classification did not show any systematic site differences; however, the instructor's role did vary by site, with Site 0 showing the highest role over years and Site 1 showing the lowest. As with the instructor classification index, no systematic trends for the subject being taught were apparent.

For the instructional focus variables, the following site trends were found. In the first instructional years, Site 0 devoted about 10% more time to decoding instruction than the average; concomitantly, about 10% less time was devoted to instruction in the meaning of sentences and texts. Site 1 devoted about 10% less time to decoding instruction with more time devoted to sentence/text instruction in instructional year 1. However, in the last two instructional years, substantially more time was devoted to decoding at this site, relative to the average, and substantially less time to instruction in word, sentence, and text meanings. In the first instructional year, Site 3 devoted large amounts of time to sentence/text instruction, with little time to word meanings, and average percents of time to decoding. In the second year, substantially more time was devoted to decoding instruction with below average amounts to word, sentence, and text meaning. Sites 2 and 5 showed the average trends with respect to these percentages of time over instructional years.

For the quality of instruction indices expressed under the instructional focus codes, Sites 1 and 5 revealed particularly low values with respect to the explicitness of the instruction on sentence/text meaning during the first two years. No other systematic site deviations from the overall trends were apparent.

Similarly, no systematic site differences were evident in the percent of time in which primary materials were used (about 90% at each of the four instructional years). Site differences did appear for the types of materials used (i.e., the degree of text usage): Site 1 showed lower than average values at instructional years 1 and 2, and Site 2 was generally lower across all years (except for year 2 which was high). For ancillary material usage, again systematic site differences were not evident. The only site differences in the quality index of these materials were at Site 2 which showed low values in the first two instructional years.

Recall that for the activity/task scales, on average, independent work accounted for about half the instructional time during the first two years, dropping to about 35% in the following two years. Conversely, group work represented about 40% of the instructional time during the first two years, increasing to about 50% for the last two years. Site 0 did not show this trend -- about a third of the time was spent in independent work and 55% in group work during each of the four instructional years. The only other site differences were found at instructional year 1, where Site 2 showed about 20% independent work and 70% group work, while Site 3 showed just the opposite (70% independent work and 25% group work).

On average, the level of formal language demand for independent work increased over the instructional years, starting at a relatively low level in the first year, and increasing to mid-level by the last two years. Sites 0, 3, and 5 showed higher values in the first two instructional years relative to those found at Sites 1 and 2. For group work, on average, formal language demand was noticeably higher than that associated with independent work, but little change in level over the instructional years occurred; site differences for these were minimal.

For the student response measures (i.e., number of nonengaged students, productivity, and noise) no systematic site differences were apparent.

In summary, the differences by site can be characterized as follows. Treating Site 0 first, this site showed the smallest average group size in each instructional year and the highest average value for the associated instructor's role. In the first instructional years, Site 0 showed a larger than average percentage of time devoted to decoding instruction, and concomitantly, a smaller percentage of time to instruction in the meaning of sentences and texts. Further, during these early grades, Site 0 spent substantially larger percentages of time in group work, and lower percentages of time in independent work (where the level of formal language demand in this independent work was relatively high).

Site 1 showed the lowest percentage of instructional time with an associated instructor and the lowest average value for the associated instructor's role. Further, this site devoted a greater than average proportion of time to sentence/text instruction in the early years (with a lower percentage to decoding instruction) -- it also showed particularly low values with respect to the explicitness of the instruction on sentence/text meaning. This was associated with both a lower than average usage of text as primary material, and relatively lower levels of formal language demand for independent work. In the latter years, a substantially larger percentage of time was devoted to decoding instruction (and a smaller percentage of time to instruction in word, sentence, and text meaning).

Site 2 generally showed the largest average group size; further, it tended to reveal the lowest average values in the degree of text usage across all years. At instructional year 1, this site showed a less than average percentage of time in independent work and an above average percentage of time in group work. Further, the level of formal language demand in independent work was relatively low.

In the first instructional year, Site 3 devoted a large percentage of time to sentence/text instruction, with a small percentage of time to word meanings, and an average percentage of time to decoding. However, in the second year, a substantially larger proportion of time was devoted to decoding instruction with below average percentages to instruction in word, sentence, and text meaning. At instructional year 1, this site showed an above average percentage of time in independent work with a below average percent of time in group work. Further, the level of formal language demand in independent work was relatively high.

During the first two instructional years, Site 5 showed particularly low values with respect to the explicitness of the instruction on sentence/text meaning. However, during these years, the level of formal language demand in independent work was relatively high.

Instruction in Spanish Reading

In this section, the Spanish reading observation summaries are presented. First, the aggregate data are discussed for the bilingual sample, then the differences in instruction found for individual sites.

Aggregate Descriptive Data

For the Spanish reading protocols, Tables 5 and 6 display the descriptive statistics (mean, standard deviation, and number of cases) for the average quality and quantity indices of instruction, respectively. These data are summarized for each of the five instructional years, and the tables are structurally identical to those displaying the English data. As in the presentation above of the English analyses, kindergarten data will not be discussed here; also the same cautions about the interpretation of the final two years of data (based on a different cohort than that represented in instructional years 1 and 2) apply here.

Before discussing these data, note that the number of students observed in Spanish reading instruction is about a third of those observed in English reading instruction -- for the first two instructional years where the entire target sample is represented, the number of students observed in English reading instruction is 215 and 224, while those observed in Spanish reading instruction is 73 and 62. This is indicative of the large exit from bilingual programs (at least an exit from any Spanish reading component that might be contained in such programs) by students at the end of kindergarten (remember, all of the target students were enrolled in bilingual programs when they were selected at kindergarten).

Turning to the descriptive data from the Spanish reading observations, first consider the number of students associated with the observed instructional groups of the target students: the trend appears to be for slightly smaller groups in the early years (around

Table 5

Reading and Mathematics Observation System - Spanish:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for the Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
NSTRMN	M	12.1	12.6	13.7	19.3	20.5
NSTRMN	S	3.9	4.3	5.2	6.4	5.4
NSTRMN	N	67	73	62	20	9
CLSSMN	M	6.3	6.5	6.7	6.7	5.9
CLSSMN	S	0.6	1.0	0.4	0.5	0.9
CLSSMN	N	67	69	62	20	9
ROLEMN	M	7.1	5.6	6.5	6.3	7.4
ROLEMN	S	1.4	1.4	1.7	1.6	0.6
ROLEMN	N	67	69	62	20	9
SBJCMN	M	7.1	7.3	7.3	6.3	3.1
SBJCMN	S	1.3	0.8	1.1	0.6	0.6
SBJCMN	N	67	73	62	20	9
IFLTMN	M	1.7	2.0	[illegible]	2.3	2.1
IFLTMN	S	0.6	0.4	0.3	0.4	0.2
IFLTMN	N	64	69	61	18	9
IFWOMN	M	2.9	2.9	2.3	2.6	3.0
IFWOMN	S	0.2	0.3	0.3	0.7	0.0
IFWOMN	N	32	28	23	15	3
FSTMN	M	1.7	1.4	1.4	1.9	1.5
FSTMN	S	0.7	0.4	0.4	0.6	0.6
FSTMN	N	24	59	54	15	9
TECHMN	M	1.1	1.2	1.4	1.5	1.6
TECHMN	S	0.2	0.2	0.3	0.4	0.3
TECHMN	N	60	46	46	20	9
LANGMN	M	.5	1.2	1.2	1.3	1.3
LANGMN	S	0.4	0.3	0.3	0.3	0.2
LANGMN	N	67	73	62	20	9
MAT1MN	M	4.3	5.3	6.0	6.3	7.3
MAT1MN	S	1.3	0.9	0.9	1.0	1.4
MAT1MN	N	67	73	62	20	9
MAT2MN	M	3.4	4.3	4.1	4.3	4.1
MAT2MN	S	.4	0.9	1.0	1.5	2.2
MAT2MN	N	62	73	60	19	9
ATNTMN	M	1.9	1.9	1.9	1.9	1.8
ATNTMN	S	0.1	0.1	0.1	0.1	0.1
ATNTMN	N	67	73	62	20	9
ATINMN	M	1.2	1.4	1.7	1.3	
ATINMN	S	0.4	0.4	0.5	0.4	
ATINMN	N	46	73	60	15	
ATLRMN	M	1.9	2.0	2.3	2.3	2.4
ATLRMN	S	0.6	0.4	0.5	0.5	0.3
ATLRMN	N	67	71	62	20	9

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
MENZMN	M	2.0	2.5	1.8	4.3	7.0
MENZMN	S	1.5	3.2	2.3	3.0	3.3
MENZMN	N	27	50	53	20	9
PROCMN	M	3.2	3.0	3.0	3.1	3.1
PROCMN	S	0.4	0.4	0.5	0.4	0.5
PROCMN	N	67	73	61	20	9
NOISMN	M	2.6	2.4	2.4	2.3	2.9
NOISMN	S	0.5	0.4	0.4	0.4	[illegible]
NOISMN	N	67	73	62	20	9
ATTIMN	M	3.4	3.5	4.0		
ATTIMN	S	0.4	0.3	.0		
ATTIMN	N	40	33	9		
CTRLMN	M	0.5	0.7	0.1	0.2	0.2
CTRLMN	S	0.6	1.2	0.3	0.3	0.4
CTRLMN	N	67	73	62	20	9
OBTMMN	M	45.3	44.7	41.1	37.2	30.9
OBTMMN	S	13.3	11.2	14.1	7.2	4.0
OBTMMN	N	67	73	62	20	9
STTMMN	M	44.5	41.2	35.7	37.2	30.9
STTMMN	S	13.2	9.7	11.2	7.2	4.1
STTMMN	N	67	73	62	20	9

Table 6

Reading and Mathematics Observation System - Spanish:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for the Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
CLSSPR	M	85.7	56.4	84.4	81.6	100.0
CLSSPR	S	18.4	28.3	20.2	22.3	0.0
CLSSPR	N	67	73	62	20	9
ROLEPR	M		62.4	85.9	82.3	100.0
ROLEPR	S	[illegible]	23.1	19.3	21.4	0.0
ROLEPR	N	67	69	62	20	9
IFLTPR	M	52.6	51.2	54.1	38.5	32.0
IFLTPR	S	24.9	21.5	23.5	29.1	18.7
IFLTPR	N	67	73	62	20	9
IFWDPR	M	11.7	2.1	9.0	21.4	19.3
IFWDPR	S	16.4	3.6	15.4	18.9	14.5
IFWDPR	N	67	73	62	20	9
IFSTPR	M	7.8	30.0	26.3	29.0	31.2
IFSTPR	S	12.7	23.5	20.2	27.6	9.3
IFSTPR	N	67	73	62	20	9
MAT1PR	M	89.7	93.6	91.8	94.4	94.7
MAT1PR	S	10.7	8.0	11.5	6.9	5.0
MAT1PR	N	67	73	62	20	9
MAT2PR	M	60.2	59.3	59.2	81.0	75.7
MAT2PR	S	23.2	19.4	26.1	19.2	15.6
MAT2PR	N	62	73	60	19	9
ATINPR	M	20.0	48.4	37.1	32.8	0.0
ATINPR	S	20.6	25.0	26.6	29.7	0.0
ATINPR	N	67	73	62	20	9
ATLRPR	M	69.2	42.8	55.7	59.9	84.7
ATLRPR	S	20.0	23.1	25.0	27.9	5.0
ATLRPR	N	67	73	62	20	9
IFTTPR	M	72.1	83.4	89.5	88.9	82.5
IFTTPR	S	22.6	13.0	15.0	10.5	5.6
IFTTPR	N	67	73	62	20	9
IFT2PR	M	19.5	32.2	35.3	50.4	50.5
IFT2PR	S	25.8	23.3	18.8	25.4	13.3
IFT2PR	N	67	73	62	20	9
ATTTPR	M	89.2	91.2	92.8	92.7	84.7
ATTTPR	S	7.8	8.2	8.4	6.1	5.0
ATTTPR	N	67	73	62	20	9
IFLTRP	M	75.8	62.0	58.2	42.3	37.8
IFLTRP	S	29.2	26.0	23.0	28.8	17.7
IFLTRP	N	67	73	62	20	9
IFWDRP	M	50.2	8.7	27.2	45.7	33.7
IFWDRP	S	37.7	14.2	37.2	39.[illegible]	25.3
IFWDRP	N	42	69	55	19	9
ATLRRP	M	77.8	47.6	60.1	65.1	100.0
ATLRRP	S	22.0	25.9	27.3	31.1	0.0
ATLRRP	N	67	73	62	20	
STTMRT	M	97.5	94.2	90.6	100.0	100.0
STTMRT	S	6.7	14.2	16.4	0.0	0.0
STTMRT	N	67	73	62	20	9

13 students, as in the English data), but much larger groups in later years (around 20 students).

For the instructor variables, for about 60% of the instruction observed, some instructor was associated with the target students in the first instructional year; this value increased to over 80% in the following years (a substantial difference from the English data). This instructor, as in the English observations, generally was a teacher, and the role played tended to be one of facilitation (somewhat lower in the first instructional year). The subject being taught generally required substantial reading, but the value suggests that other subjects were also observed.

For the instructional focus variables, about 90% of the observed instruction contained instructional foci. About half the time was devoted to decoding instruction during the first and second instructional years, falling to about 30% by the fourth year, mirroring the English observations. This instruction tended to be focused on isolated units in kindergarten, then on non-explicit letter-sound pairings for the remaining grades -- again, the relatively low variability in this measure suggests little explicit letter-sound work. The amount of time spent on word meanings was negligible during the first year (only 2%), increasing to about 10% in the second year, and to 20% in the last two years. The instruction that was offered tended to be quite explicit (as it was in the English data). About 30% of the instructional focus time was on instruction in the meaning of sentences and texts, remaining at this level for all four years. The quality of this instruction was fairly stable across the years, and was generally non-explicit.

For materials, the use and quality of the primary materials closely resembled that found in the English reading observations (employed about 90% of the time over the four years with an increase toward more textual material). Ancillary materials were used about 60% of the time during the first two instructional years, and about 75% for years 3 and 4 (somewhat higher here than in English). As in the English observations, the quality index of these materials was stable over the four instructional years, and was lower than the primary materials.

For the activity/task scales, about 90% of the observed time contained instructional activity/tasks. Independent work accounted for about half the instructional time during the first year, dropping to about 35% in the following two years, and to zero in the final year (but remember that the number of cases is small here). Conversely, group work represented about 40% of the instructional time during the first year, increasing to about 60% for the next two years, and to 85% for the final year. For independent work, the level of formal language demand increased over the instructional years, starting at a relatively low level in the first year, and increasing to mid-level by the last year. For group work, however, little change in level of formal language demand occurred (perhaps lower in the first year); it was, as in the English data, noticeably higher than that associated with independent work.

For the student response measures, the number of nonengaged students was low, but substantially higher in the latter two years; productivity was rated as medium in each year, while noise tended to be low. Finally, the number of momentary controls was low in all years but the first, where it averaged about .7 per observation.

Site Differences

Given this overall description of the Spanish reading observation data, individual site differences are discussed next. The quality and quantity summary measures for each site appear in Appendix F, tabled in the same format represented in Tables 5 and 6. Again, the site differences discussed have not been subjected to statistical test, and are given only as supplemental information to the site descriptions provided in Volume 2. Given the small number of cases at instructional years 3 and 4 within each site, differences for these later years will not be discussed.

For the number of students associated with the observed instructional groups of the target students, each of the five sites showed the general trend for slightly smaller groups in the early years and somewhat larger groups in later years. Site 3 showed the smallest average group sizes with about 10 students per group in the first two instructional years.

Concerning the instructor variables, Site 0 showed the highest percentage of instructional time with an associated instructor -- about 95% over the four instructional years. Sites 1, 3, and 5 showed the lowest percentage at year 1 (about 40%), with Site 3 continuing to show a low value in the second instructional year (about 60%). The associated instructor's classification did not show any substantial site differences; however the instructor's role did vary by site. Site 1 generally showed the lowest average role value over years, while Site 2 showed the highest value in the first year, followed in the second year by the lowest average value. For the subject being taught, noticeably low values were apparent at Site 0 for the first year, and Site 2 for the second year, indicating that many non-reading subjects were taught during the scheduled reading period at these two sites.

For the instructional focus variables, the following site trends were found with respect to the quantity indices. In the first instructional year, Sites 0, 1, and 5 devoted about half of the instructional time to decoding instruction, while Sites 2 and 3 provided decoding instruction for only 20% of the time. In the second instructional year, decoding instruction was given for about half the observed time, slightly less in Site 2 (about 40%), but substantially more in Site 5 (about 65%). All sites gave little attention to word meaning instruction in the first two instructional years (less than 10%), except for Site 0, where such instruction represented about 25%

of the observed time during the second instructional year. Instruction in the meaning of sentences and texts represented about 30% of the instructional focus time in the aggregate over the first two instructional years. For instructional year 1, Site 0 devoted about 15% to such instruction, Site 1 about 45%, and Site 3 about 75%. In year 2, Site 0 remained at 15%, while Site 2 dropped to about 33%, and Site 3 to about 40%.

For the quality of instruction indices expressed under the instructional focus codes, decoding instruction tended to have a mid-scale average (i.e., non-explicit instruction), with Site 1 being somewhat lower than the others. Instruction in word meaning tended to have a high value, but again, Site 1 was slightly lower. Note that the quality indices under instruction in word meaning are missing for some instructional years under Sites 2 and 3 (in Tables 5 and 7 of Appendix F) -- as seen in the corresponding tabled quantity values (Tables 6 and 8 of Appendix F), such instruction was never observed at these sites for these years. Finally, for instruction in the meaning of sentences and texts, aggregate values tended to be low -- the highest average value found was mid-scale for Site 0 during instructional year 1.

Few site differences were evident in the percent of time in which primary materials were used (about 90% of the observed instruction in the aggregate); Site 0 was about 10% below the average value at year 1, while Site 2 was about 20% below the average in the second instructional year. Site differences for the types of primary materials used were also minimal. For ancillary material usage, low use (around 30% when compared to the average value of 60%) was found at Site 0 for the first instructional year, and at Sites 2 and 3 (also around 80%) for the second instructional year. No substantial site differences in the quality index of these materials were evident.

Recall that for the activity/task scales, on average, independent work accounted for about half the instructional time during the first year, dropping to about 35% in the following year. Conversely, group work represented about 40% of the instructional time during the first, increasing to about 60% for the second year. Site 0 did not show this trend -- about 15% of the time was spend in independent work and 65% in group work during the first two years. The only other site differences were found at Site 2, which showed about 30% independent work and 60% group work for instructional year 1, with about 45% in both for instructional year 2; Site 3 showed the opposite trend in the first instructional year (70% independent work and 30% group work).

On average, the level of formal language demand for independent work increased over the instructional years, starting at a relatively low level in the first year, and increasing to mid-level by the last two years. Site differences in this trend were minimal, with the exception of Site 1 which showed a value noticeably lower than the average for the second instructional year. For group work, on average, formal language demand was noticeably higher than that associated with independent work, but little change in level over the instructional years occurred -- site differences with respect to these were minimal.

For the student response measures, the number of nonengaged students was substantially higher at Site 2, especially during the first instructional year. The rated noise level was noticeably higher at Sites 0, 1, and 2, relative to that found at Sites 3 and 5. Finally, productivity was rated highest at Site 0.

The differences by site can be summarized as follows. Site 0 showed the highest percentage of instructional time with an associated instructor, but low values for the amount of reading required by the subject being taught. Decoding instruction accounted for about half the instructional time during the initial years (the average). Instruction in sentence and text meaning represented about 15% of the instructional time (as compared to an average value of 30%) -- the quality of this instruction was the highest found during the first year. Instruction in word meaning was negligible during the first year, but quite large during the second year. For the activity/task scales, this site spent about 15% of the instructional time in independent work and 65% in group work. Finally, the noise level was higher at this site, but productivity was also rated high.

Site 1 was among the sites with the lowest percentage of instructional time with an associated instructor and the lowest average value for the associated instructor's role. About half of the instructional time was devoted to decoding instruction, and most of the remaining time to instruction in the meaning of sentences and texts -- little time was devoted to word meaning. The quality of decoding instruction was somewhat lower in this site than at the others, as was that for the limited amount of instruction in sentence/text meaning. For independent work during the second instructional year, which represented about 40% of the instructional time, the level of formal language demand was noticeably low. The rated noise level was among the highest of the sites.

Site 2 had the highest instructor role value for the first instructional year, but the lowest for the following year. Low values were also found for the subject being taught during the second instructional year. Decoding instruction constituted only 20% of the instructional focus during the first year, increasing to 40% in the subsequent year. Little time was devoted to word meanings (indeed, no time was noted during the last three years); instruction in the meaning of sentences and texts represented almost half the instruction in the first year, falling to about a third in the second year. Use of both primary and ancillary materials was about 25% below average in the second instructional year. Independent work represented 30% of the instruction in the first year, with group work representing about 60% -- in the second year both accounted for about 45% of the instructional time. The number of nonengaged students was substantially higher at Site 2, and the rated noise level was also among the highest values found.

Site 3 revealed the smallest average group size for the first two instructional years. This site also showed the lowest percent of time with an instructor for the first two years (about 40% and 60%, respectively). The percentage of time devoted to decoding instruction was below average in the first year (at about 20%), but average in the second year (at 50%); no time on word meanings was noted, but instruction in the meaning of sentences and texts was above average in both instructional years, substantially so in the first year. For independent and group work, Site 3 showed substantially more independent work (at about 70%) and less group work (representing about 30%) in the first year, but followed the trend in the second year of about 40% independent work and 60% group work. The rated noise level was among the lowest values found.

Finally, for Site 5, the only noticeable difference from the aggregate picture concerned the percent of time devoted to decoding in the second year (65%), which was above the average value of about 50%.

In closing, in this section the descriptive statistics for the RAMOS summary indices for the bilingual sample have been described, treating English and Spanish reading instruction separately. Discussions of the trends with respect to individual sites have also been included. Given this basis, the next section describes the procedures and results of a further reduction of these instructional indices.

Factor Analyses

The 21 instructional quality indices and the 16 quantity indices described earlier, represent a great reduction in data from the original protocols. Nevertheless, they still represent a sizeable number of indices. In order to further reduce these instructional data, factor analysis techniques were employed. Such analyses were conducted over the entire set of individual protocol summaries obtained from the bilingual sample (254 students), disregarding student, grade, and date of observation, but treating English and Spanish reading observations separately (1293 and 347 individual student summaries, respectively, obtained over the five years of data collection).

An examination of the descriptive statistics and correlations based on the individual protocol summary measures argued for the deletion of some of the original set of variables from the factor analyses for the following reasons: (1) negligible variance, (2) strong correlations with other variables entered in the analysis, and (3) an insufficient number of cases for which a scale value was available (any value which was based on less than 80% of the total number of cases). Thus, of the set of 37 defined quality and quantity indices, only 25 were entered in the English analysis, 22 in the Spanish analysis. The specific variables deleted and the reasons for their deletion are provided below.

For the English analysis, only one variable was deleted due to negligible variance. This was NSTDPR, the percentage of time during which the number of students contained in an observed group was coded, which was always 100%. A number of variables were deleted due to their strong correlations with other entered variables (deleted variables are listed first):

(1) ROLEPR, the percentage of time during which an associated instructor's role was coded, as expected, was highly correlated (.96) with CLSSPR, the percentage of time during which an instructor was associated with the group,

(2) IFWDPR, the percentage of time devoted to instruction in word meaning, was correlated (.77) with IFWDRP, the percentage of time devoted to word meaning instruction relative to the percentage of time devoted to non-decoding instruction (i.e., instruction in word, sentence, and text meaning),

(3) ATLRPR, the percentage of time devoted to listening and responding in groups, was negatively correlated (-.92), as expected, with ATINPR, the percentage of time spent in independent work,

(4) IFT2PR, the percentage of time devoted to non-decoding instruction, was highly correlated (.83) with IFSTPR, the percentage of time devoted to instruction in sentence and text meaning, which could be expected given the small amount of time devoted to instruction in word meanings,

(5) ATTTPR, the total average percentage of time devoted to instructional activity/tasks, was highly correlated (.90) with ATNTMN, which is also an index of the amount of time devoted to instructional activity/tasks,

(6) IFLTRP, the percentage of time devoted to decoding instruction relative to the percentage of time devoted to instruction (disregarding non-instructional time), was highly correlated (.87) with IFLTPR, the percentage of time devoted to decoding instruction (considering all instruction observed), and

(7) ATLRRP, the percentage of time devoted to group instruction relative to the total activity time, was negatively correlated (-.98), as expected, with ATINPR, the percentage of time devoted to independent work.

Finally, a number of variables were deleted due to a relatively small number of cases for which a value was available (any value which was based on less than 80% of the total number of cases). These included the following: (1) IFWDMN (652 cases or 5?% of the sample), again reflecting the little attention given to instruction in word meanings, and (2) ATT1MN and ATT1PR, the attention rating variables associated with the initial RAMOS version employed during the first two data collection years (236 cases, representing 18% of the sample).

A similar pattern of deletions held for the Spanish variables. As in the English data, the only variable deleted due to negligible variance was NSTDPR, which was always 100%. The same set of variables deleted in the English data due to their strong correlations with other entered variables were also deleted here: (1) ROLEPR was highly correlated (.97) with CLSSPR, (2) IFWDPR was correlated (.81) with IFWDRP, (3) ATLRPR was negatively correlated (-.93) with ATINPR, (4) IFT2PR was highly correlated (.83) with IFSTPR, (5) ATTTPR was highly correlated (.94) with ATNTMN, (6) IFLTRP was highly correlated (.88) with IFLTPR, and (7) ATLRRP was negatively correlated (-.97) with ATINPR. Finally, a number of variables were deleted due the relatively small number of cases for which a value was available, more here than found in the English analyses. These included the following: (1) IFWDMN (125 cases or 36% of the sample), (2) IFSTMN (245 cases or 71%), (3) TECHMN and TECHPR (252 cases or 73%), (3) NEN2MN and NEN2PR (237 cases or 69%), and (4) ATT1MN and ATT1PR (110 cases or 32%).

English Factor Analysis

In this section, the results of the factor analysis conducted on the English individual student summaries are presented, followed by descriptive data on the subsequent computation of factor scores for the individual student yearly summaries.

Derivation of Factors

In the analysis, eight factors had eigenvalues above 1, but only the initial seven factors were employed in subsequent analyses. The first factor accounted for 13.8% of the variance, with the next successive six factors accounting for 10.3%, 8.5%, 7.4%, 6.3%, 5.6%, and 4.9%, respectively, for a total of 56.7% explained variance (chance expecta on is 28%). Of the 25 variables entered in the analysis, each l ded on at least one factor, all loading on a single factor except fo ATINPR which loaded on two factors. Table 7 gives the factor loadings for the variables entered in the English analysis. The rows of the table define the entered variables, and the columns define the factors, using mnemonic names which best reflect the instructional dimensions identified in the analysis. Only variables with loadings greater than .45 were used in the subsequent computation of factor scores, and thus, only these variables are discussed below.

The first factor gives a strong positive weight to an emphasis on reading (SBJCMN), to the allocation of observed time to instruction (reflected in both ATNTMN and ATNTPR), and to a high proportion of use of the primary materials (MAT1PR). The negative weight to NEN2MN indicates a strong relation with the number of engaged (as opposed to nonengaged) students. The secondary variables, showing weaker, but not negligible loadings, include (1) a positive weight to the instruc-

Table 7

Reading and Mathematics Observation System - English:
Varimax Rotated Factor Matrix
Based on 1293 Observations from the Bilingual Sample

		Factor						
		ETT	DGI	QFL	ADC	PRD	SMT	NST
	NSTDMN	-0.096	-0.045	0.049	0.004	-0.107	-0.106	0.778
	CLSSMN	0.049	-0.166	0.016	-0.238	0.465	0.126	0.285
	ROLEMN	0.386	0.367	0.127	0.060	0.151	-0.181	-0.308
	SBJCMN	0.524	-0.031	0.417	0.116	-0.042	0.053	-0.199
	IFLTMN	0.037	-0.050	0.595	-0.059	-0.047	-0.019	0.184
	IFSTMN	0.048	0.315	-0.107	0.120	0.570	0.062	-0.012
	TECHMN	-0.258	0.105	0.270	-0.057	0.331	0.596	-0.044
	LANGMN	0.006	-0.447	0.429	0.032	0.084	-0.246	-0.170
	MAT1MN	-0.031	-0.115	0.671	-0.059	-0.103	0.166	-0.105
	MAT2MN	0.073	-0.530	0.239	-0.004	-0.125	-0.079	-0.052
	ATNTMN	0.764	0.124	0.090	-0.046	-0.106	-0.130	0.106
	ATINMN	-0.080	0.059	0.608	0.248	0.378	-0.007	0.099
RAMOS	ATLRMN	0.197	0.101	0.594	-0.179	-0.184	-0.169	0.063
Variable	VEN2MN	-0.642	0.241	0.152	0.014	-0.180	0.274	0.282
	PRDCMN	-0.017	0.099	-0.075	0.072	0.689	-0.029	-0.310
	NOISMN	-0.356	0.526	0.020	0.103	-0.305	0.369	-0.219
	CTRLMN	0.073	0.032	-0.254	0.147	-0.173	0.607	0.102
	CLSSPR	-0.039	0.743	0.158	0.145	0.085	-0.323	-0.010
	IFLTPR	0.251	-0.019	0.056	0.774	-0.117	-0.126	-0.129
	IFSTPR	0.204	-0.085	0.138	-0.906	-0.064	0.059	[illegible]
	MAT1PR	0.646	-0.162	0.002	-0.080	-0.012	0.141	-0.064
	MAT2PR	0.081	-0.258	0.067	-0.188	0.176	0.634	-0.140
	ATNTPR	0.478	0.001	0.178	0.041	0.037	0.163	0.389
	ATINPR	0.297	-0.576	0.027	-0.142	-0.448	0.041	0.124
	IFWDRP	-0.113	0.212	-0.037	0.532	0.244	0.218	0.096

6·2

tor's role (ROLEMN), indicating a trend toward more direct instruction, and (2) a negative weight to noise (NOISMN), meaning relatively quieter instructional environments. As the variable loadings suggest a measure of the engaged reading time with text, the factor has been named ETT, for engaged text time.

The second factor is more complex. It shows strong positive weights for the percent of time an instructor is involved in the instructional activity (CLSSPR), and for the role the instructor is playing (ROLEMN), an index of direct instruction. The negative weight to ATINPR indicates relatively fewer independent activity/task assignments (work is being largely carried out in groups), and thus, the noise level is higher (positive weight to NOISMN). The remaining two variables of LANGMN and MAT2MN, both with negative loadings, are more difficult to interpret. They suggest that the higher the factor score, the greater the use of Spanish (in these English reading protocols), and the lower the quality of the ancillary materials. Although the label is oversimplified, the factor seems to provide an index of direct group instruction, labelled DGI.

The third factor consists entirely of quality indices, all with positive loadings. The strong weight for MAT1MN indicates a trend toward more work involving text; the loading for IFLTMN shows that when the instructional focus is at the level of letter-sound units (which happened half the time), the focus is toward analytic strategies rather than whole word recognition or oral reading. Finally, the loadings for ATINMN and ATLRMN indicate higher levels of formal language demand, whether instruction is carried out independently or in groups. Secondary variables include SBJCMN and LANGMN, both with positive weights, indicating an emphasis on reading and on the use of English. In general, the factor may best be described as an index of the quality of formal language instruction, or QFL.

The variables loading on the fourth factor all concern the emphasis on the amount of decoding. The strong positive loading for IFLTPR indicates a greater proportion of time spent on letter-sound units than at the level of word, sentence, or text. The strong negative weight to IFSTPR, the percentage of time devoted to instruction at the sentence and text levels, is expected given its inverse relationship with IFLTPR. The final variable loading on this factor is IFWDRP, the proportion of time spent at the word level relative to the sentence/text level. Its positive loading suggests that the teacher who focuses on decoding does little sentence/text work, presumably because students are weak in decoding. There are no secondary variable loadings. In general, the factor seems to represent the amount of decoding instruction, or ADC.

The highest loading variable on the fifth factor is the positive loading for PRDCMN, the rating of productivity. It is associated with the quality of work when the focus is on the sentence/text level, as indicated by the positive loading of IFSTMN. The final two variables, which show weak relationships with the factor, are for CLSSMN and ATINPR. The first, with a positive loading, indicates that higher levels of productivity are associated with teacher directed instruction as opposed to that directed by other individuals (e.g., aides). The negative loading for ATINPR, the proportion of independent work, supports the contention of increased productivity when instruction is teacher-directed. Overall, the factor best describes the conditions that promote high individual productivity, labelled PRD, although the inclusion of IFSTMN suggests this is an oversimplification.

The sixth factor is quite puzzling. It shows three positive loadings: (1) MAT2PR, the percentage of time involving the use of secondary materials, (2) CTRLMN, the number of momentary controls, and (3) TECHMN, the binary index of part-to-whole versus whole-to-part instructional strategies. It has been labelled SMT, for use of secondary materials, since this variable shows the largest loading (but not by much) -- it is clearly oversimplified.

The final factor shows a single primary loading, NSTDMN, the number of students. Secondary variables are (1) ROLEMN, a negative loading, indicating that the instructor's role tends more toward facilitation, less toward direct instruction, (2) PRDCMN, a negative loading showing decreasing productivity, and (3) ATNTPR, a positive loading, indicating an increase in the percentage of time in which activity/task codes are applicable (indicative of fewer transitional activities). In general, the factor seems to represent effects associated with large groups (a less direct mode of instruction, reduced productivity, and fewer major shifts in activities) -- it is simply labelled number of students, or NST.

In summary, the seven factors identified in the analysis are (1) engaged text time, an index of reading time where students are engaged with text materials, (2) direct group instruction, an index of direct instruction which is aimed at groups, rather than individual students, (3) the quality of formal language, an index of the formal language demands made upon the students, (4) the amount of decoding instruction, (5) student productivity, (6) the use of secondary materials, though the interpretation of this factor is not straightforward, and (7) the number of students constituting an instructional group.

Descriptive Statistics on Factor Scores

On the basis of this analysis, factor scores were created by first transforming the yearly averaged values (those final summary indices described earlier which were the averaged Fall and Spring average scale values) to z-scores, and then weighting the relevant component z-scores for each factor (those showing loadings greater than .45) by the appropriate loading values.

In Tables 8 and 9, the descriptive statistics for the computed English factor scores for the bilingual sample are presented for each site under each instructional year (Table 8 containing instructional year 0 statistics, and Table 9 displaying instructional years 1 through 4).

Table 8

Reading and Mathematics Observation System - English:
Descriptive Statistics on Factor Scores Overall and by Site
for Instructional Year 0 for the Bilingual Sample

Factor	Statistic	Overall	Site 0	Site 1	Site 2	Site 3	Site 5
ETT	M	0.008	0.145	0.216	0.045	0.195	-0.211
	S	0.349	0.147	0.184	0.217	0.423	0.292
	N	140	30	6	13	35	56
OGI	M	-0.007	0.223	0.166	0.191	0.006	-0.205
	S	0.323	0.187	0.213	0.165	0.226	0.349
	N	140	30	6	13	35	56
QFL	M	0.000	0.103	-0.283	-0.242	0.160	-0.067
	S	0.366	0.272	0.206	0.210	0.457	0.32[illegible]
	N	140	30	6	13	35	5[illegible]
ADC	M	0.060	-0.225	0.468	0.090	0.293	0.018
	S	0.567	0.350	0.408	0.489	0.730	0.503
	N	140	30	6	13	35	56
PRD	M	-0.017	0.078	-0.153	-0.4[illegible]	0.358	-0.177
	S	0.377	0.190	0.171	0.377	0.207	0.321
	N	140	30	6	13	35	56
SMT	M	-0.032	0.777	-0.031	-0.035	-0.259	-0.324
	S	0.499	0.381	0.388	0.230	0.187	0.171
	N	140	30	6	13	35	56
NST	M	0.000	-0.396	-0.005	0.145	-0.973	0.786
	S	0.999	0.209	0.497	0.612	0.456	0.973
	N	140	30	6	13	35	56

Table 9

Reading and Mathematics Observation System - English:
Descriptive Statistics on Factor Scores Overall and by Site
for Instructional Years 1-4 for the Bilingual Sample

INSTRUCTIONAL YEAR 1

Factor	Statistic	Overall	Site 0	Site 1	Site 2	Site 3	Site 5
ETT	M	-0.001	-0.116	-0.111	-0.033	0.249	-0.189
	S	0.338	0.294	0.193	0.131	0.191	0.434
	N	215	53	12	28	71	51
OGI	M	0.000	0.259	-0.200	0.390	-0.182	-0.183
	S	0.361	0.290	0.430	0.316	0.158	0.304
	N	215	53	12	28	71	51
OFL	M	-0.013	0.197	-0.376	-0.600	0.027	0.118
	S	0.433	0.323	0.435	0.406	0.252	0.430
	N	215	53	12	28	71	51
ADC	M	0.022	0.361	-0.204	0.366	-0.430	0.165
	S	0.645	0.378	0.907	0.628	0.343	0.751
	N	215	53	12	28	71	51
PRD	M	0.000	0.309	-0.198	0.238	-0.230	-0.087
	S	0.353	0.286	0.293	0.281	0.190	0.331
	N	215	53	12	28	71	51
SMT	M	0.009	0.293	0.378	0.086	-0.286	-0.002
	S	0.442	0.220	0.257	0.596	0.171	0.543
	N	215	53	12	28	71	51
NST	M	0.000	-0.827	0.038	1.665	-0.223	0.246
	S	1.000	0.568	0.662	0.397	0.579	0.938
	N	215	53	12	28	71	51

INSTRUCTIONAL YEAR 2

Factor	Statistic	Overall	Site 0	Site 1	Site 2	Site 3	Site 5
ETT	M	-0.003	-0.255	-0.364	-0.055	0.132	0.148
	S	0.348	0.438	0.586	0.231	0.221	0.169
	N	224	50	11	34	76	53
OGI	M	-0.001	0.279	-0.214	-0.095	-0.023	-0.130
	S	0.335	0.313	0.346	0.375	0.183	0.344
	N	224	50	11	34	76	53
OFL	M	-0.002	-0.008	-0.619	-0.046	0.205	-0.137
	S	0.380	0.410	0.214	0.365	0.183	0.394
	N	224	50	11	34	76	53
ADC	M	0.010	0.121	-0.221	-0.124	0.133	-0.138
	S	0.619	0.555	0.517	0.652	0.428	0.831
	N	224	50	11	34	76	53
PRD	M	-0.001	0.413	-0.115	-0.043	-0.176	-0.072
	S	0.329	0.276	0.234	0.316	0.198	0.200
	N	224	50	11	34	76	53
SMT	M	0.008	0.191	0.610	0.440	-0.414	0.042
	S	0.489	0.326	0.287	0.411	0.235	0.481
	N	224	50	11	34	76	53
NST	M	0.000	-1.158	0.249	0.140	0.584	0.114
	S	0.999	0.569	0.589	0.663	0.618	1.120
	N	224	50	11	34	76	53

INSTRUCTIONAL YEAR 3

Factor	Statistic	Overall	Site 0	Site 1	Site 2
ETT	M	0.000	0.011	-0.135	0.044
	S	0.349	0.361	0.401	0.298
	N	94	50	14	30
OGI	M	0.000	0.102	-0.105	-0.121
	S	0.370	0.237	0.306	0.513
	N	94	50	14	30
OFL	M	-0.002	0.105	-0.085	-0.142
	S	0.427	0.426	0.327	0.431
	N	94	50	14	30
ADC	M	0.000	-0.071	0.335	-0.038
	S	0.618	0.518	0.197	0.829
	N	94	50	14	30
PRD	M	0.000	0.019	-0.214	0.067
	S	0.274	0.271	0.220	0.260
	N	94	50	14	30
SMT	M	0.000	-0.080	0.197	0.041
	S	0.367	0.353	0.371	0.359
	N	94	50	14	30
NST	M	0.000	-0.415	0.876	0.282
	S	1.000	0.902	0.784	0.898
	N	94	50	14	30

INSTRUCTIONAL YEAR 4

Factor	Statistic	Overall	Site 0	Site 1	Site 2
ETT	M	0.001	0.057	-0.108	-0.051
	S	0.303	0.279	0.345	0.319
	N	59	33	9	17
OGI	M	0.000	-0.006	-0.115	0.071
	S	0.333	0.301	0.295	0.405
	N	59	33	9	17
OFL	M	-0.002	0.042	0.031	-0.105
	S	0.350	0.339	0.355	0.368
	N	59	33	9	17
ADC	M	0.006	-0.048	0.417	-0.106
	S	0.625	0.410	0.944	0.780
	N	59	33	9	17
PRD	M	-0.001	0.004	-0.032	0.004
	S	0.377	0.478	0.193	0.198
	N	59	33	9	17
SMT	M	0.000	0.000	0.008	-0.005
	S	0.265	0.244	0.341	0.279
	N	59	33	9	17
NST	M	0.000	-0.212	0.116	0.350
	S	0.999	0.897	0.922	1.165
	N	59	33	9	17

606

Note that the procedure for deriving factor scores involves standardization of the averaged summary indices first, then weighting and summing of factor component values. Thus, the resulting factor scores are not expected to conform to a standard distribution with a mean of 0 and a standard deviation of 1. As seen in Table 9, the mean of each of the (overall) factor scores at each instructional year is close to 0 (ranging from -0.013 to 0.022 over the four instructional years); the standard deviations, however, show a much larger range (from 0.265 to 1.000). Although specific site differences in factor score averages are apparent, they will not be discussed here, as they largely reflect differences in the averaged summary indices that were discussed earlier.

In Tables 10 through 14, the correlations for the bilingual sample between the computed English factor scores and each yearly averaged summary z-score (regardless of whether or not the variable was included in the factor analysis) are presented, for instructional years 0 through 4, respectively. These correlations are discussed below, again, treating only those for instructional years 1 through 4.

For the first factor, engaged text time (ETT), moderate positive relations are apparent at each instructional year (about .55) for the total percentage of instructional focus time devoted to literacy (IFTTPR). This variable was not included in the factor analysis, and its moderate correlation here suggests that greater ETT values are associated with greater amounts of time devoted to literacy instruction. The ETT factor shows substantial positive relations at each instructional year (about .7) with the total percentage of time devoted to instructional activity/tasks (ATTTPR). This variable was not included in the factor analysis due to its high correlation with ATNTMN, which did load on this factor -- both are indices of the amount of time devoted to instructional activity/tasks. Although the quality of the primary materials, MAT1MN, did not show a substantial loading here, its correlations with the factor over the four years are not negligible, ranging from .2 to .5. These correlations are, however, consistent with the interpretation given the factor. For variables found to load on the ETT factor, two show interesting patterns. First, the yearly correlations with ATNTPR are low (about .35), reflecting its relatively low weighting in the factor analysis solution. Second, the correlations for the percentage of primary material usage, MAT1PR, are higher in the first year (.76) than in any of the succeeding years (about .45).

For the second factor, direct group instruction (DGI), ROLEPR, which was not included in the factor analysis, reveals a high correlation (about .8), as expected given its strong relationship to CLSSPR which did load on this factor. Similarly, both ATLRPR and ATLRRP, neither of which were included in the factor analysis, correlate highly (about .8) with the factor -- again, such is expected given the high negative correlations between these variables and ATINPR, which did load on this factor and shows negative correlations with it over the four instructional years. For variables loading on the DGI factor, two show noteworthy patterns. First, the contribution of

Table 10

Reading and Mathematics Observation System - English:
Correlations between Scale 7-Scores and Factor Scores
for Instructional Year 0 for the Bilingual Sample

Scale	Entered in Analysis	Factor Scores (N=140) ETT	DGI	QFL	ADC	PRD	SMT	NST
NSTDMN	yes	-0.326	-0.131	-0.275	-0.136	-0.158	-0.155	1.000
CLSSMN	yes	0.071	-0.094	0.137	-0.160	0.611	-0.101	0.103
ROLEMN	yes	0.518	0.550	0.399	0.029	0.297	0.391	-0.517
SBJCMN	yes	0.584	0.285	0.262	0.283	0.449	0.211	-0.355
IFLTMN	yes	-0.064	0.238	0.556	-0.105	0.232	-0.054	0.209
IFWDMN	no	0.468	0.338	0.341	0.201	-0.156	0.199	-0.840
IFSTMN	yes	-0.170	0.585	0.061	0.053	0.467	0.330	0.205
TECHMN	yes	0.178	0.447	0.149	-0.269	0.070	0.389	-0.241
LANGMN	yes	-0.024	-0.377	0.058	-0.052	0.284	-0.487	-0.134
MAT1MN	yes	0.410	0.155	0.589	0.083	-0.054	0.018	-0.261
MAT2MN	yes	0.077	-0.509	-0.155	0.232	0.021	-0.219	-0.250
ATNTMN	yes	0.724	0.089	-0.048	-0.099	0.073	0.109	0.019
ATINMN	yes	-0.120	0.117	0.559	-0.036	0.245	0.297	-0.106
ATLRMN	yes	0.424	0.344	0.569	0.059	0.271	0.060	-0.383
NEN2MN	yes	-0.631	-0.208	-0.130	-0.245	-0.434	-0.291	0.459
PRDCMN	yes	0.633	0.248	0.304	0.142	0.803	0.094	-0.450
NOISMN	yes	-0.061	0.554	0.108	-0.065	-0.121	0.409	0.187
ATT1MN	no	0.024	-0.218	0.433	-0.492	0.802	0.530	-0.319
CTRLMN	yes	0.107	0.428	0.096	-0.191	0.022	0.815	-0.020
OBTMMN	no	-0.361	-0.267	0.124	-0.085	-0.038	-0.078	0.338
STTMMN	no	-0.071	-0.268	0.086	-0.130	-0.081	-0.128	0.338
CLSSPR	yes	0.254	0.793	0.328	0.129	0.365	0.162	-0.192
ROLEPR	no	0.242	0.780	0.323	0.112	0.358	0.156	-0.192
IFLTPR	yes	0.452	0.271	0.273	0.707	0.434	0.012	-0.431
IFWDPR	no	0.306	0.489	0.140	-0.147	-0.101	0.487	-0.297
IFSTPR	yes	0.351	0.316	0.224	-0.738	0.319	0.414	-0.287
MAT1PR	yes	0.301	0.155	-0.007	-0.009	0.382	0.356	-0.304
MAT2PR	yes	0.080	0.306	-0.053	-0.055	0.178	0.798	-0.161
ATNTPR	yes	0.566	-0.070	0.020	-0.155	-0.093	0.024	-0.005
ATINPR	yes	0.057	-0.461	-0.198	0.059	-0.715	-0.043	0.094
ATLRPR	no	0.402	0.445	0.198	-0.083	0.659	0.107	-0.139
IFTTPR	no	0.646	0.557	0.375	0.299	0.441	0.364	-0.602
IFT2PR	no	0.399	0.507	0.217	-0.497	0.101	0.558	-0.359
ATTTPR	no	0.753	0.107	0.057	-0.057	0.113	0.119	-0.103
IFLTRP	no	-0.176	-0.401	-0.072	0.623	-0.051	-0.432	0.133
IFWDRP	yes	0.363	0.386	-0.049	0.740	-0.268	0.288	-0.383
ATLRRP	no	0.099	0.520	0.229	-0.067	0.741	0.088	-0.095
STTMRT	no	0.537	-0.055	-0.057	-0.118	-0.101	-0.121	0.044

Table 11

Reading and Mathematics Observation System - English:
Correlations between Scale Z-Scores and Factor Scores
for Instructional Year 1 for the Bilingual Sample

Scale	Entered in Analysis	Factor Scores (N=215) ETT	DGI	QFL	ADC	PRD	SMT	NST
NSTDMN	yes	-0.210	0.075	-0.369	0.091	-0.039	0.043	1.000
CLSSMN	yes	-0.021	-0.079	-0.039	-0.174	0.384	0.073	-0.125
ROLEMN	yes	0.305	0.632	-0.034	0.305	0.381	0.146	-0.382
SBJCMN	yes	0.651	0.084	-0.050	-0.089	-0.108	-0.130	-0.070
IFLTMN	yes	-0.088	-0.117	0.661	-0.095	0.146	-0.166	-0.266
IFWDMN	no	0.010	0.179	0.375	0.079	0.313	0.242	-0.324
IFSTMN	yes	0.182	0.325	0.228	0.177	0.656	0.165	-0.187
TECHMN	yes	-0.440	[illegible]	-0.002	0.438	0.427	0.621	-0.215
LANGMN	yes	0.201	[illegible]	0.275	0.060	0.010	-0.031	-0.162
MAT1MN	yes	0.206	0.173	0.720	-0.023	0.116	-0.047	-0.321
MAT2MN	yes	-0.189	-0.639	-0.121	-0.205	-0.307	0.074	-0.115
ATNTMN	yes	0.753	-0.151	-0.278	-0.230	-0.243	-0.094	-0.012
ATINMN	yes	0.002	0.268	0.726	0.230	0.280	-0.057	-0.224
ATLRMN	yes	0.096	-0.477	0.603	-0.287	-0.402	-0.312	-0.137
NEK2MN	yes	-0.795	0.210	0.081	0.263	0.287	-0.061	0.352
PRDCMN	yes	-0.278	0.377	0.030	0.305	0.772	0.330	0.093
NOISMN	yes	-0.190	0.568	0.028	0.429	0.293	0.542	-0.190
ATTIMN	no	-0.145	0.462	0.259	0.141	0.822	-0.294	0.086
CTRLMN	yes	-0.088	0.119	-0.336	0.357	0.094	0.726	0.302
OBTMMN	no	-0.188	0.141	0.558	0.189	0.133	0.016	-0.451
STTMMN	no	0.053	-0.269	0.332	-0.113	-0.097	-0.157	-0.228
CLSSPR	yes	0.055	0.783	0.013	0.287	0.340	-0.135	0.340
ROLEPR	no	0.116	0.747	-0.033	0.420	0.322	-0.065	0.313
IFLTPR	yes	0.082	0.329	-0.102	0.819	0.173	0.192	-0.036
IFWDPR	no	-0.109	0.462	0.058	0.304	0.275	0.086	0.056
IFSTPR	yes	0.303	-0.439	-0.093	-0.932	-0.335	-0.255	-0.061
MAT1PR	yes	0.756	-0.057	-0.271	-0.184	-0.106	-0.047	-0.024
MAT2PR	yes	0.059	-0.209	-0.027	-0.112	0.059	0.665	0.171
ATNTPR	yes	0.414	-0.010	-0.140	0.024	-0.068	-0.093	0.150
ATINPR	yes	0.327	-0.782	0.054	-0.584	-0.689	-0.149	-0.078
ATLRPR	no	-0.108	0.802	-0.145	0.561	0.662	0.130	0.091
IFTTPR	no	0.504	0.106	-0.249	-0.030	-0.080	-0.051	-0.108
IFT2PR	no	0.274	-0.261	-0.073	-0.858	-0.234	-0.233	-0.039
ATTTPR	no	0.779	-0.152	-0.269	-0.233	-0.273	-0.097	0.019
IFLTRP	no	-0.171	0.299	0.035	0.875	0.204	0.189	0.013
IFWDRP	yes	-0.254	0.556	0.015	0.690	0.366	0.241	0.177
ATLRRP	no	-0.210	0.812	-0.072	0.567	0.677	0.125	0.067
STTMRT	no	0.308	-0.419	-0.300	-0.324	-0.248	-0.161	0.287

Table 12

Reading and Mathematics Observation System - English:
Correlations between Scale Z-Scores and Factor Scores
for Instructional Year 2 for the Bilingual Sample

Scale	Entered in Analysis	Factor Scores (N=224) ETT	DGI	QFL	ADC	PRD	SMT	NST
NSTDMN	yes	0.141	-0.250	0.181	0.014	-0.494	-0.114	1.000
CLSSMN	yes	-0.162	-0.081	-0.287	-0.298	0.378	0.299	0.005
ROLEMN	yes	0.264	0.471	0.291	0.027	0.207	-0.420	-0.159
SBJCMN	yes	0.688	0.039	0.236	0.076	-0.032	-0.210	-0.026
IFLTMN	yes	0.172	0.189	0.561	0.141	-0.002	-0.148	0.131
IFWDMN	no	-0.015	0.202	-0.286	-0.022	0.189	-0.040	-0.303
IFSTMN	yes	-0.075	0.184	0.115	0.159	0.640	-0.124	-0.284
TECHMN	yes	-0.397	0.247	-0.178	-0.174	0.342	0.763	-0.342
LANGMN	yes	0.007	-0.336	-0.079	0.257	0.135	-0.030	-0.115
MAT1MN	yes	0.338	0.218	0.538	-0.112	-0.123	0.038	-0.037
MAT2MN	yes	0.153	-0.424	0.005	0.035	-0.227	-0.241	0.076
ATNTMN	yes	0.783	-0.043	0.235	-0.110	-0.248	-0.279	0.246
ATINMN	yes	0.152	0.123	0.712	0.342	0.020	-0.371	0.141
ATLRMN	yes	0.224	0.289	0.638	-0.078	-0.239	-0.417	0.256
NEW2MN	yes	-0.569	0.345	0.038	-0.111	0.033	0.255	0.041
PRDCMN	yes	-0.130	0.106	-0.190	0.087	0.749	0.230	-0.459
NOISMN	yes	-0.428	0.571	0.110	-0.039	0.192	0.164	-0.261
ATT1MN	no	0.337	-0.230	0.264	-0.576	0.837	-0.136	-0.410
CTRLMN	yes	-0.165	-0.245	-0.320	-0.258	-0.040	0.802	0.186
OBTMMN	no	0.046	0.203	0.308	0.264	0.048	-0.521	0.009
STTMMN	no	0.248	0.055	0.249	0.175	-0.032	-0.457	0.165
CLSSPR	yes	0.115	0.818	0.474	0.105	0.191	-0.391	-0.088
ROLEPR	no	0.108	0.762	0.383	0.066	0.204	-0.298	-0.126
IFLTPR	yes	0.182	-0.057	0.217	0.838	-0.123	-0.289	0.096
IFWDPR	no	-0.053	-0.035	-0.034	0.129	0.134	-0.027	-0.031
IFSTPR	yes	0.125	0.027	-0.129	-0.953	-0.006	0.248	-0.009
MAT1PR	yes	0.313	-0.308	-0.117	-0.020	-0.291	-0.057	0.153
MAT2PR	yes	-0.188	-0.245	-0.324	-0.213	0.176	0.797	-0.195
ATNTPR	yes	0.334	-0.340	-0.077	0.119	-0.321	-0.131	0.211
ATINPR	yes	0.325	-0.753	-0.109	0.074	-0.566	-0.003	0.406
ATLRPR	no	0.017	0.760	0.243	-0.107	0.428	-0.138	-0.302
IFTTPR	no	0.557	-0.098	0.144	-0.062	-0.122	-0.120	0.142
IFT2PR	no	0.092	0.008	-0.140	-0.841	0.030	0.222	-0.024
ATTTPR	no	0.771	-0.081	0.271	-0.061	-0.365	-0.302	0.273
IFLTRP	no	-0.042	-0.041	0.150	0.852	-0.085	-0.264	0.036
IFWDRP	yes	-0.108	0.011	-0.082	0.561	0.144	-0.112	-0.098
ATLRRP	no	-0.161	0.771	0.178	-0.081	0.519	-0.072	-0.393
STTMRT	no	0.384	-0.205	-0.051	-0.090	-0.118	0.091	0.270

Table 13

Reading and Mathematics Observation System - English:
Correlations between Scale Z-Scores and Factor Scores
for Instructional Year 3 for the Bilingual Sample

		Factor Scores (N=94)						
Scale	Entered in Analysis	ETT	DGI	QFL	ADC	PRD	SMT	NST
NSTDMN	yes	-0.181	-0.008	-0.018	-0.109	0.026	-0.045	1.000
CLSSMN	yes	-0.051	-0.215	-0.421	-0.171	0.417	0.112	0.281
ROLEMN	yes	0.284	0.737	0.386	0.143	0.176	-0.191	-0.293
SBJCMN	yes	0.685	0.222	0.395	0.347	0.161	-0.038	-0.113
IFLTMN	yes	0.211	-0.061	0.633	0.016	-0.051	-0.125	-0.106
IFWDMN	no	0.059	0.224	-0.202	-0.002	0.047	-0.006	0.313
IFSTMN	yes	0.062	-0.104	0.099	[illegible]	0.691	0.024	-0.041
TECHMN	yes	-0.016	-0.169	-0.132	0.136	-0.099	0.572	0.028
LANGMN	yes	0.025	-0.410	0.359	-0.438	0.242	0.022	-0.124
MAT1MN	yes	0.524	0.110	0.720	-0.012	-0.043	0.125	0.021
MAT2MN	yes	0.200	-0.439	-0.094	0.249	-0.507	0.059	-0.056
ATNTMN	yes	0.575	0.247	0.036	0.243	0.179	-0.295	-0.234
ATINMN	yes	0.238	0.350	0.655	-0.008	0.412	-0.137	0.054
ATLRMN	yes	0.047	0.128	0.737	-0.202	0.142	-0.079	0.008
NEN2MN	yes	-0.655	0.343	-0.030	0.121	-0.210	-0.093	0.157
PRDCMN	yes	0.147	-0.077	0.499	-0.398	0.554	-0.101	-0.167
NOISMN	yes	-0.081	0.593	0.308	0.400	-0.317	0.080	-0.115
ATT1MN	no	-	-	-	-	-	-	-
CTRLMN	yes	-0.120	0.201	0.044	0.204	0.110	0.703	0.035
OBTMMN	no	0.091	0.193	0.386	0.017	0.110	-0.125	-0.302
STTMMN	no	0.159	0.182	0.361	0.030	0.132	-0.099	-0.304
CLSSPR	yes	0.086	0.846	0.239	0.297	0.008	-0.109	0.105
ROLEPR	no	0.093	0.792	0.265	0.253	0.019	-0.088	0.152
IFLTPR	yes	0.257	0.206	-0.078	0.818	-0.175	0.096	-0.223
IFWDPR	no	0.060	0.194	0.134	0.260	-0.193	-0.004	0.227
IFSTPR	yes	-0.038	-0.290	0.032	-0.964	0.349	-0.096	0.059
MAT1PR	yes	0.576	-0.398	-0.062	0.000	0.027	0.136	0.100
MAT2PR	yes	0.061	-0.214	-0.026	-0.146	-0.072	0.528	-0.138
ATNTPR	yes	0.397	-0.115	-0.038	0.058	-0.149	0.069	0.015
ATINPR	yes	-0.107	-0.770	0.032	-0.254	-0.287	0.074	-0.076
ATLRPR	no	0.290	0.768	-0.021	0.304	0.295	-0.155	-0.044
IFTTPR	no	0.594	0.141	0.102	0.163	0.059	0.000	-0.014
IFT2PR	no	0.005	-0.145	0.124	-0.750	0.202	-0.096	0.217
ATTTPR	no	0.573	0.326	0.020	0.255	0.145	-0.263	-0.313
IFLTRP	no	0.143	0.160	-0.088	0.784	-0.194	0.118	-0.198
IFWDRP	yes	0.177	0.339	-0.041	0.651	-0.289	0.053	0.046
ATLRRP	no	0.226	0.770	0.000	0.284	0.301	-0.104	-0.015
STTMRT	no	0.242	-0.023	-0.094	0.023	0.107	0.078	-0.043

Table 14

Reading and Mathematics Observation System - English:
Correlations Between Scale Z-Scores and Factor Scores
for Instructional Year 4 for the Bilingual Sample

		Factor Scores (N=59)						
Scale	Entered in Analysis	ETT	DGI	QFL	ADC	PRD	SMT	NST
NSTDMN	yes	-0.135	0.458	0.078	-0.019	0.330	-0.219	1.000
CLSSMN	yes	-0.018	0.098	-0.377	0.162	0.636	-0.218	0.329
ROLEMN	yes	0.251	0.677	0.342	0.119	0.334	-0.134	0.139
SBJCMN	yes	0.486	0.049	0.020	-0.177	-0.112	-0.051	-0.179
IFLTMN	yes	0.336	0.036	0.486	-0.028	0.079	0.154	0.005
IFWDMN	no	-0.097	0.214	0.077	-0.155	-0.241	0.064	0.053
IFSTMN	yes	-0.006	0.279	-0.281	0.036	0.773	-0.243	0.271
TECHMN	yes	0.024	0.121	0.052	0.343	0.283	0.152	0.154
LANGMN	yes	-0.053	-0.503	0.232	-0.413	-0.100	0.465	-0.245
MAT1MN	yes	0.473	-0.075	0.688	0.084	-0.306	0.048	-0.009
MAT2MN	yes	-0.107	-0.711	-0.040	-0.173	-0.547	0.308	-0.368
ATNTMN	yes	0.656	0.179	0.184	0.393	0.252	-0.026	-0.065
ATINMN	yes	-0.100	0.333	0.550	0.248	0.065	-0.048	0.469
ATLRMN	yes	0.050	0.019	0.497	-0.002	-0.449	0.290	-0.205
NEN2MN	yes	-0.484	0.071	0.040	0.000	-0.363	-0.081	0.105
PRDCMN	yes	0.210	0.070	-0.085	0.050	0.783	-0.122	0.083
NOISMN	yes	-0.277	-0.190	0.068	-0.431	-0.691	0.189	-0.132
ATT1MN	no	-	-	-	-	-	-	-
CTRLMN	yes	0.117	-0.189	0.305	-0.170	-0.662	0.604	-0.244
OBTMMN	no	-0.024	-0.292	0.126	-0.082	-0.304	0.274	-0.292
STTMMN	no	0.085	-0.219	0.048	-0.195	-0.210	0.204	-0.362
CLSSPR	yes	-0.021	0.882	0.201	0.257	0.463	-0.465	0.533
ROLEPR	no	-0.067	0.765	0.114	0.290	0.461	-0.338	0.504
IFLTPR	yes	0.243	0.422	0.091	0.798	0.262	-0.416	0.135
IFWDPR	no	0.112	-0.249	0.007	0.062	-0.120	0.084	-0.278
IFSTPR	yes	-0.036	-0.078	-0.177	-0.963	-0.050	0.246	0.081
MAT1PR	yes	0.437	-0.177	-0.132	-0.240	-0.153	0.190	-0.034
MAT2PR	yes	-0.055	-0.578	-0.084	-0.540	-0.120	0.532	-0.186
ATNTPR	yes	0.292	-0.069	0.033	-0.135	-0.078	0.126	0.025
ATINPR	yes	0.026	-0.794	0.032	-0.062	-0.525	0.581	-0.337
ATLRPR	no	0.265	0.818	0.043	0.196	0.589	-0.539	0.305
IFTTPR	no	0.491	0.294	-0.145	-0.236	0.221	-0.186	0.045
IFT2PR	no	0.035	-0.221	-0.154	-0.823	-0.118	0.271	-0.096
ATTTPR	no	0.629	0.172	0.159	0.300	0.216	0.002	-0.019
IFLTRP	no	0.056	0.262	0.108	0.821	0.171	-0.340	0.168
IFWDRP	yes	0.117	-0.120	0.144	0.694	-0.093	-0.063	-0.209
ATLRRP	no	0.051	0.793	-0.026	0.052	0.551	-0.600	0.336
STTMRT	no	0.191	0.104	-0.136	-0.237	0.147	-0.095	-0.156

612

LANGMN increases in magnitude over the instructional years (from about -.25 to -.50), suggesting that the trend for the increased use of Spanish that this variable represents in the direct group instruction factor, becomes more pronounced in later grade levels. Second, the positive, moderate contribution of NOISMN to this factor is fairly stable for the first three instructional years, but effectively drops to zero in the final instructional year.

The third factor, quality of formal language (QFL), reveals an interesting correlation trend for three of the four loading variables. For IFLTMN, ATINMN, and ATLRMN, the correlations are all greater in the first three instructional years (averaging about .65 for each variable) than in the fourth year (averaging about .50), thus requiring some caution in the interpretation of this factor with respect to the last year.

For the factor reflecting the amount of decoding (ADC), two variables which were not included in the factor analysis show strong correlations with the factor over the four instructional years, namely, IFT2PR (about -.8) and IFLTRP (about .8). Both are expected, given their strong relationships to IFLTPR which did load on this factor. Two variables which did not load on this factor show interesting correlations. First, LANGMN is negatively correlated with the factor in instructional years 3 and 4 (at about -.4), suggesting that the use of Spanish increases in the latter years as the amount of decoding instruction increases. Second, NOISMN shows a positive correlation with the factor in years 1 and 3 (of about .4), but a negative relation of the same magnitude in year 4 -- these suggest that the noise level of the instructional group increases in the early grades with increases in the amount of decoding time, but decreases with such increases in the latter grades.

Recall that the next factor, productivity (PRD), was complex, with the highest loading for PRDCMN, and weaker, but not insignificant, loadings for CLSSMN, IFSTMN, and ATINPR. The correlation pattern of this factor with the yearly averaged summary z-scores reflects this structure. First, CLSSMN shows positive relationships at each instructional year. The lower coefficients found in the first three years (about .4) relative to the last year (.64) indicate the relatively weaker contribution of this variable to the factor in the early grades. The variables ATLRPR and ATLRRP, neither of which were included in the factor analysis, show the expected positive relationships to the factor score given their negative relationships to the negatively loading variable ATINPR. Interestingly, for MAT2MN, which did not load on this factor, negative correlations are found at each instructional year, larger in the latter two years (about -.5) than in the initial two years (about -.3). This suggests that the quality of the secondary materials is generally negatively related to the factor, but more strongly so in the latter grade levels. Finally, the ATT1MN variable, which was not included in the factor analysis, shows strong positive relationships with the factor for the years in which it was employed (about .8 in years 1 and 2). Such a relationship is consistent with the productivity interpretation given this factor (productivtivity increases with increases in student attention to the instructional task).

The sixth factor, secondary materials (SMT), was also a complex factor, showing loadings for three variables, TECHMN, CTRLMN, and MAT2PR. Its correlation pattern shows a single interesting perturbation: TECHMN, the dichotomous index of whether instruction was largely from parts-to-whole or from whole-to-parts, is strongly related to the factor in the first three years (about .65), but only weakly related in the fourth year (.15). Indeed, in the fourth year, six other (non-loading) variables show coefficients greater than .4, thus requiring caution in the interpretation of this factor for this final year.

The final factor, number of students (NST), consisted of a single variable, namely the average number of students associated with the target student's instructional group. From the correlation tables, only two other variables show consist relationships over the four instructional years, namely, the length of the observation (OBTMMN) and the amount of time the target student was observed during the observation (STTMMN). Both of these variables show small negative relationships with the factor score (about -.3) in each of the instructional years except for the second. This suggests that smaller group sizes tended to be observed during longer observations. Such would be expected if small group work represented a significant part of a given teacher's reading instruction period, for if classroom size is constant, and the teacher works with each reading group, then working in relatively smaller groups requires longer periods, and thus longer observation lengths.

Spanish Factor Analysis

In this section, the results of the factor analysis conducted on the Spanish individual student summaries are presented, followed by descriptive data on the subsequent computation of factor scores.

Derivation of Factors

In this analysis as in the English one, again eight factors had eigenvalues above 1, but only the initial seven factors were employed in subsequent analyses. The first factor accounted for 14.7% of the variance, with the next successive six factors accounting for 11.7%, 10.5%, 7.8%, 7.4%, 6.2%, and 5.4%, respectively, for a total of 63.8% explained variance (chance expectation is 32%). Of the 22 variables entered in the analysis, each loaded on at least one factor except for ATNTPR; of these all loaded on a single factor with the exception of ROLEMN which loaded on three factors. Table 15 gives the factor loadings for the variables entered in the Spanish analysis. As in the English table, the rows define the entered variables, and the columns define the factors, employing mnemonic names which best reflect the instructional dimensions identified in the analysis. As before, only variables with loadings greater than .45 were used in the subsequent

Table 15

Reading and Mathematics Observation System - Spanish:
Varimax Rotated Factor Matrix
based on 347 Observations from the Bilingual Sample

		Factor						
		QFL	DGI	ETT	NST	ADC	SMT	CNT
	NSTDMN	0.180	0.158	-0.222	0.688	-0.079	0.182	0.096
	CLSSMN	0.467	-0.172	-0.066	0.184	-0.051	-0.208	0.388
	ROLEMN	0.065	0.512	0.464	-0.454	-0.122	0.018	-0.201
	SBJCMN	0.376	0.055	0.631	-0.006	-0.001	0.286	0.122
	IFLTMN	0.707	-0.034	-0.091	-0.177	-0.064	-0.073	0.292
	LANGMN	-0.220	0.311	-0.027	0.255	-0.106	-0.372	-0.497
	MAT1MN	0.650	-0.019	0.041	0.097	0.068	0.111	-0.208
	MAT2MN	-0.188	0.085	0.077	0.233	-0.233	0.579	0.177
	ATNTMN	-0.064	-0.238	0.811	-0.001	-0.036	-0.090	-0.027
	ATINMN	0.640	0.002	0.071	0.014	0.246	-0.101	-0.068
RAMOS	ATLRMN	0.630	0.045	-0.028	-0.062	-0.400	0.221	-0.052
Variable	PRDCMN	0.149	0.133	-0.075	-0.819	0.056	0.095	0.061
	NOISMN	-0.279	0.598	-0.255	0.203	0.113	0.184	0.074
	CTRLMN	-0.120	0.130	-0.015	0.011	0.094	-0.054	0.638
	CLSSPR	0.135	0.836	-0.073	0.066	0.111	-0.028	-0.106
	IFLTPR	0.102	0.012	0.299	0.079	0.795	0.078	0.281
	IFSTPR	0.001	-0.214	0.180	0.195	-0.848	0.084	0.099
	MAT1PR	-0.118	-0.130	0.306	-0.050	0.094	0.175	-0.005
	MAT2PR	0.014	-0.217	0.077	-0.214	0.131	0.098	-0.223
	ATNTPR	0.236	0.047	0.218	0.289	-0.010	0.398	-0.102
	ATINPR	0.079	-0.770	0.149	0.219	-0.076	0.225	-0.135
	IFWDRP	-0.061	0.337	-0.167	-0.362	0.425	0.138	-0 510

615

computation of factor scores, and thus, only these variables are discussed below.

The first factor shows loadings exclusively for quality variables, all with positive weights. The strongest weight is toward increased analytic, high-level decoding instruction (IFLTMN). The weight for MAT1MN shows a trend toward basals whenever primary materials are employed. The weights for ATINMN and ATLRMN indicate that whether work is taking place independently or in groups, the factor reflects a higher level of formal language demand. The weakest weight is for the instructor classification variable (CLSSMN), which tends to be the teacher whenever an instructor is present (about three-fourths of the time). The only secondary variable is for SBJCMN, indicating a trend toward increased reliance on reading. This factor strongly resembles that of the third factor found in the English analysis, and is likewise, labelled quality of formal language, or QFL.

The strongest loading found for the second factor is for the percent of time that an instructor is present (CLSSPR), with a smaller positive weight to the instructor role (ROLEMN). The large negative weight to ATINPR indicates that instruction tends to be conducted in groups rather than independently. Finally, the positive weighting of NOISMN shows that the instruction tends to be more "noisy" (probably because more group instruction is taking place). Secondary variables include positive weights to LANGMN and IFWDRP. In general, this factor suggests direct group instruction delivered by an instructor. It resembles the second factor found in the English analysis, and is likewise, labelled direct group instruction, or DGI.

The third factor shows large positive weights to the allocation of observed time to instructional activities (ATNTMN), and to a high percentage of usage of primary materials (MAT1PR). Smaller positive weights are given to an emphasis on reading instruction, and to a trend toward more direct instruction (SBJCMN and ROLEMN, respectively). No secondary variables appear. The factor resembles the first factor found in the English analysis, and is also labelled engaged text time, or ETT.

The fourth factor reveals a strong negative weighting for rated productivity (PRDCMN) and a strong positive weighting for the number of students (NSTDMN). Also the negative loading of ROLEMN indicates that when an instructor is present, the role tends to be away from direct instruction and toward facilitation. In general, the structure suggests large group activities that are lower in both productivity and the role played by the instructor. Although it is more complex than the last factor identified in the English analysis, it is also named number of students, or NST.

The fifth factor has large loadings for two quantity indices, one a positive loading for IFLTPR, and the other a negative loading for IFSTPR -- both reflect the percentage of time devoted to decoding instruction. Resembling the fourth factor of the English analysis, this factor constitutes an index of the amount of decoding, or ADC.

The sixth factor largely reflects the use of ancillary materials, showing positive weights for both the quality of such materials (MAT2MN), and the percent of time in which they are used (MAT2PR). Secondary variables include a negative weight to LANGMN, suggesting that more Spanish is associated with the use of secondary materials in these Spanish reading observations, and a positive weight to ATNTPR, indicating an increase in the percent of time when activity/task codes are applicable. Given the large weights for the secondary materials, this factor is labelled secondary materials, or SMT.

The final factor is complex. Its strongest weighting is with the number of momentary controls (CTRLMN), which is inversely related to the language of instruction (LANGMN) -- it suggests that more such management interruptions are found when the language of instruction tends more toward the exclusive use of Spanish. The negative weight to IFWDRP suggests that such controls are more prevalent as the time spent on word meanings increases relative to instruction on meaning represented at higher level structures (sentences and texts). This factor has been labelled control (CNT), although this is clearly an oversimplification.

In summary, the seven factors identified in the Spanish factor analysis are (1) quality of formal language (the third English factor derived), (2) direct group instruction (corresponding to the second English factor), (3) engaged text time (the first English factor), (4) number of students (the last English factor), (5) amount of decoding (the fourth English factor), (6) secondary material usage (no directly corresponding English factor) and (7) control (a complex factor also without an English correspondence).

Descriptive Statistics on Factor Scores

On the basis of these analyses, factor scores were created following the same procedure used in the computation of factor scores for the English summaries: first transforming the yearly averaged values to z-scores, and then weighting the relevant component z-scores for each factor (those showing loadings greater than .45) by the appropriate loading values.

In Tables 16 and 17, the descriptive statistics for the computed Spanish factor scores for the bilingual sample are presented for each site under each instructional year (Table 16 containing statistics for instructional year 0, and Table 17 displaying instructional years 1 through 4).

Again, the resulting factor scores are not expected to conform to a standard distribution, given the procedure followed in their derivation. As seen in Table 17, the mean of each of the (overall) factor scores at each instructional year is close to 0 (ranging from -0.012 to 0.015 over the four instructional years). As in the English data, the standard deviations show a much larger range (from 0.103 to 0.783). Although specific site differences in factor score averages are apparent, they will not be discussed here, as they largely reflect

Table 16

Reading and Mathematics Observation System - Spanish:
Descriptive Statistics on Factor Scores Overall and by Site
for Instructional Year 0 for the Bilingual Sample

Factor	Statistic	Overall	Site 0	Site 1	Site 2	Site 3	Site 5
QFL	M	0.005	0.021	-0.125	-0.390		0.218
	S	0.310	0.164	0.182	0.142		0.279
	N	67	20	7	13		27
DGI	M	0.000	0.306	-0.204	0.005		-0.177
	S	0.419	0.364	0.177	0.328		0.421
	N	67	20	7	13		27
ETT	M	0.002	0.354	0.207	-0.023		-0.299
	S	0.452	0.253	0.137	0.390		0.442
	N	67	20	7	13		27
NST	M	0.000	-0.403	0.345	0.156		0.134
	S	0.471	0.330	0.319	0.514		0.386
	N	67	20	7	13		27
ADC	M	0.000	-0.347	-0.021	0.196		0.169
	S	0.724	1.143	0.484	0.452		0.292
	N	67	20	7	13		27
SMT	M	-0.001	0.400	-0.155	0.082		-0.346
	S	0.460	0.365	0.254	0.302		0.355
	N	62	20	7	12		23
[illegible]	M	-0.006	0.278	0.010	-0.107		-0.173
	S	0.413	0.363	0.446	0.265		0.402
	N	67	20	7	13		27

618

Table 17

Reading and Mathematics Observation System - Spanish:
Descriptive Statistics on Factor Scores Overall and by Site
for Instructional Years 1-4 for the Bilingual Sample

INSTRUCTIONAL YEAR 1

Factor	Statistic	Overall	Site 0	Site 1	Site 2	Site 3	Site 5
QFL	M	-0.012	0.027	-0.143	-0.419	-0.137	0.213
	S	0.390	0.429	0.264	0.123	0.307	0.337
	N	73	9	8	15	7	34
DGI	M	-0.012	0.613	-0.226	0.522	-0.304	-0.303
	S	0.530	0.013	0.319	0.355	0.281	0.412
	N	73	9	8	15	7	34
ETT	M	0.005	-0.976	0.187	0.363	0.522	-0.041
	S	0.567	0.404	0.276	0.252	0.106	0.478
	N	73	9	8	15	7	34
VST	M	-0.008	-0.503	0.262	0.244	-0.154	-0.021
	S	0.455	0.429	0.303	0.640	0.146	0.302
	N	73	9	8	15	7	34
ADC	M	0.000	0.172	-0.223	0.268	-1.368	0.170
	S	0.783	0.376	0.511	0.549	0.552	0.764
	N	73	9	8	15	7	34
SMT	M	0.000	-0.395	0.477	0.081	-0.090	-0.026
	S	0.498	0.617	0.264	0.270	0.350	0.527
	N	73	9	8	15	7	34
CNT	M	-0.004	0.531	0.198	-0.371	0.105	-0.053
	S	0.392	0.473	0.050	0.106	0.120	0.344
	N	73	9	8	15	7	34

INSTRUCTIONAL YEAR 2

Factor	Statistic	Overall	Site 0	Site 1	Site 2	Site 3	Site 5
QFL	M	-0.001	0.190	-0.685	-0.088	0.260	0.105
	S	0.377	0.223	0.242	0.179	0.252	0.307
	N	62	15	8	12	5	22
DGI	M	0.000	0.454	-0.230	-0.123	-0.276	-0.096
	S	0.433	0.122	0.323	0.482	0.375	0.385
	N	62	15	8	12	5	22
ETT	M	0.000	-0.070	-0.066	-0.656	0.348	0.351
	S	0.523	0.231	0.376	0.694	0.055	0.223
	N	62	15	8	12	5	22
VST	M	-0.004	-0.596	0.297	0.544	-0.042	0.000
	S	0.534	0.093	0.398	0.531	0.150	0.401
	N	62	15	8	12	5	22
ADC	M	-0.001	0.157	-0.093	-0.489	-0.420	0.285
	S	0.729	0.411	0.650	0.539	0.525	0.895
	N	62	15	8	12	5	22
SMT	M	0.000	-0.016	0.310	-0.239	-0.300	0.074
	S	0.461	0.274	0.329	0.272	0.616	0.561
	N	60	15	8	10	5	22
CNT	M	0.015	-0.224	0.196	-0.032	0.042	0.131
	S	0.364	0.134	0.202	0.128	0.176	0.524
	N	62	15	8	12	5	22

INSTRUCTIONAL YEAR 3

Factor	Statistic	Overall	Site 0	Site 1	Site 2
QFL	M	0.000	0.114	-0.174	-0.170
	S	0.240	0.173	0.305	0.017
	N	20	12	5	3
DGI	M	0.000	0.217	-0.736	0.360
	S	0.547	0.259	0.424	0.536
	N	20	12	5	3
ETT	M	0.001	0.057	0.030	-0.270
	S	0.373	0.403	0.252	0.415
	N	20	12	5	3
VST	M	0.000	-0.217	0.156	0.606
	S	0.486	0.439	0.369	0.109
	N	20	12	5	3
ADC	M	0.000	-0.152	0.236	0.216
	S	0.760	0.772	0.708	0.929
	N	20	12	5	3
SMT	M	0.000	-0.194	0.420	0.120
	S	0.418	0.395	0.134	0.000
	N	19	12	5	2
CNT	M	0.010	-0.037	-0.128	0.433
	S	0.253	0.152	0.243	0.179
	N	20	12	5	3

INSTRUCTIONAL YEAR 4

Factor	Statistic	Overall	Site 0	Site 1	Site 2
QFL	M	-0.002	0.017		-0.070
	S	0.119	0.130		0.000
	N	9	7		2
DGI	M	0.001	0.038		-0.130
	S	0.103	0.083		0.000
	N	9	7		2
ETT	M	-0.002	0.045		-0.170
	S	0.548	0.623		0.000
	N	9	7		2
VST	M	0.001	-0.238		0.840
	S	0.584	0.393		0.000
	N	9	7		2
ADC	M	0.000	0.180		-0.630
	S	0.668	0.652		0.000
	N	9	7		2
SMT	M	0.002	0.020		-0.060
	S	0.119	0.132		0.000
	N	9	7		2
CNT	M	-0.001	-0.104		0.360
	S	0.289	0.235		0.000
	N	9	7		2

differences in the averaged summary indices that were discussed earlier.

In Tables 18 through 22, the correlations for the bilingual sample between the computed Spanish factor scores and each yearly averaged summary z-score (regardless of whether or not the variable was included in the factor analysis) are presented, for instructional years 0 through 4, respectively.

Before discussing the correlation trends, note that the coefficients for instructional year 4 (Table 22) are uniformly high. Of the nine target students represented in this fourth instructional year, six come from the same classroom. Further, only a single Spanish reading observation was obtained from that classroom during the year, and all six students were assigned to the same group during that observation, showing little individual instructional variations. Similarly, two of the remaining three students came from a single classroom, again observed only once delivering Spanish reading instruction during the year, with little individual differences between the instruction received by the two students. As such, these data show the high interrelations found in Table 22, and they will not be considered in the discussions below of the correlational pattern over instructional years. A similar account can be given for the relatively high coefficients found within the third instructional year, though the problem is not as severe given the larger sample size and greater representation of classrooms.

For the first factor, quality of formal language (QFL), the lowest weighting was for CLSSMN, and its correlations with the factor are generally low (indeed, slightly negative in the third instructional year). ROLEMN, which did not load on this factor, shows a small positive correlation (about .4) in all years but the first, indicating that the quality of formal language increases as the instructor's role tends toward direct instruction. Finally, OBTMMN and STTMMN, both indicators of the length of observation (but not entered in the analysis), show moderate positive correlations with the factor in all three years (about .5), suggesting that greater quality of formal language tended to be observed in longer observations.

For direct group instruction (DGI), NEN2MN, which was not included in the factor analysis, shows a moderate positive correlation (about .5) in the first and third instructional years. This indicates that as group instruction increases so does the number of students not engaged in the activity assigned. Both ATLRPR and ATLRRP, neither of which were included in the factor analysis, correlate highly (about .8) with the factor -- as in the English data, such is expected given the high negative correlations between these variables and ATINPR, which did load on this factor and shows negative correlations with it over the instructional years.

For the third factor, engaged text time (ETT), NEN2MN, which was not included in the factor analysis, shows a moderate negative correlation (about -.6) for all years, supporting the interpretation of

620

Table 18

Reading and Mathematics Observation System - Spanish:
Correlations between Scale Z-Scores and Factor Scores
for Instructional Year 0 for the Bilingual Sample

Scale	Entered in Analysis	Factor Scores (N=67) QFL	DGI	ETT	NST	AOC	SMT	CNT
NSTDMN	yes	0.263	0.256	-0.008	0.651	0.196	-0.518	0.118
CLSSMN	yes	0.620	-0.097	-0.104	0.199	0.014	-0.496	0.287
ROLEMN	yes	0.052	0.556	0.480	-0.703	-0.187	0.433	0.140
SBJCMN	yes	0.218	0.537	0.514	0.016	0.235	0.179	0.458
IFLTMN	yes	0.640	0.046	-0.102	-0.184	0.305	-0.202	0.153
IFWDMN	no	0.306	-0.015	0.098	-0.259	-0.281	0.236	0.191
IFSTMN	no	-0.205	0.302	0.007	-0.566	-0.793	0.495	0.017
TECHMN	no	-0.243	0.067	0.309	-0.224	-0.710	0.371	-0.002
LANGMN	yes	-0.073	-0.238	-0.448	0.079	-0.482	-0.038	-0.723
MAT1MN	yes	0.260	0.097	-0.032	0.407	0.214	-0.281	0.191
MAT2MN	yes	-0.254	0.320	0.332	-0.341	-0.020	0.653	0.128
ATNTMN	yes	-0.328	-0.072	0.749	0.113	0.011	0.234	0.318
ATINMN	yes	0.522	0.018	-0.453	0.246	0.101	-0.435	-0.322
ATLRMN	yes	0.538	0.148	-0.005	-0.477	-0.229	0.264	-0.003
NEN2MN	no	0.084	-0.532	-0.099	0.157	0.179	-0.161	0.344
PRDCMN	yes	0.138	0.314	0.172	-0.789	-0.265	0.226	0.153
NOISMN	yes	-0.324	0.562	0.281	-0.007	-0.168	0.350	0.229
ATT1MN	no	0.307	0.343	0.095	-0.725	-0.422	0.659	0.135
CTRLMN	yes	-0.040	0.394	0.305	0.003	-0.114	0.204	0.695
OBTMMN	no	0.443	-0.222	-0.303	-0.221	-0.058	0.025	-0.106
STTMMN	no	0.395	-0.196	-0.246	-0.268	-0.065	0.038	-0.127
CLSSPR	yes	0.203	0.686	0.338	0.196	0.275	-0.113	0.313
ROLEPR	no	0.203	0.686	0.338	0.196	0.275	-0.113	0.313
IFLTPR	yes	0.163	0.191	0.316	0.213	0.872	0.086	0.484
IFWDPR	no	-0.470	0.205	0.225	-0.317	-0.614	0.471	-0.191
IFSTPR	yes	-0.162	0.118	0.281	-0.328	-0.889	0.379	0.134
MAT1PR	yes	-0.373	0.267	0.807	-0.300	-0.070	0.498	0.401
MAT2PR	yes	-0.289	-0.007	0.377	-0.393	-0.213	0.778	0.032
ATNTPR	yes	-0.037	0.376	0.255	-0.168	-0.076	0.325	0.090
ATINPR	yes	-0.253	-0.631	-0.007	0.255	-0.034	0.009	-0.360
ATLRPR	no	0.147	0.707	0.297	-0.273	0.020	0.153	0.477
IFTTPR	no	-0.253	0.427	0.6[illegible]1	-0.181	0.014	0.742	0.470
IFT2PR	no	-0.379	0.188	0.282	-0.364	-0.829	0.488	-0.055
ATTTPR	no	-0.289	0.156	0.745	-0.029	-0.037	0.425	0.279
IFLTRP	no	0.405	-0.172	-0.248	0.345	0.839	-0.368	0.009
IFWDRP	yes	-0.823	-0.011	-0.066	0.074	0.116	0.362	-0.684
ATLRRP	no	0.215	0.662	0.078	-0.239	0.026	0.005	0.389
STTMRT	no	-0.093	0.119	0.248	-0.217	-0.008	0.059	-0.097

Table 19

Reading and Mathematics Observation System - Spanish:
Correlations between Scale Z-Scores and Factor Scores
for Instructional Year 1 for the Bilingual Sample

Scale	Entered in Analysis	Factor Scores (N=73) QFL	DGI	ETT	NST	ADC	SMT	CNT
NSTDMN	yes	0.021	0.412	-0.363	0.803	0.357	-0.106	-0.082
CLSSMN	yes	0.383	-0.120	-0.280	0.189	0.096	-0.424	0.296
ROLEMN	yes	-0.118	0.604	0.477	-0.364	-0.109	-0.037	-0.221
SBJCMN	yes	-0.167	-0.335	0.853	-0.022	-0.318	0.259	-0.195
IFLTMN	yes	0.786	0.013	-0.431	-0.056	0.116	-0.266	0.300
IFWDMN	no	-0.174	0.007	-0.213	-0.279	0.209	-0.123	-0.399
IFSTMN	no	0.098	0.391	-0.260	-0.479	0.155	-0.285	0.355
TECHMN	no	-0.123	0.605	0.050	0.260	0.439	-0.162	-0.191
LANGMN	yes	-0.131	0.381	0.263	0.202	0.244	0.095	-0.773
MAT1MN	yes	0.656	-0.145	-0.220	-0.316	0.214	-0.102	0.331
MAT2MN	yes	-0.152	0.452	0.079	0.023	-0.121	0.728	-0.064
ATNTMN	yes	-0.286	-0.374	0.911	-0.086	-0.259	0.317	-0.198
ATINMN	yes	0.562	0.045	0.050	-0.116	0.455	-0.023	0.108
ATLRMN	yes	0.654	-0.337	-0.150	0.017	-0.026	-0.255	-0.106
NEN2MN	no	0.027	0.595	-0.493	0.892	0.494	-0.302	-0.338
PRDCMN	yes	0.259	-0.025	-0.396	-0.771	-0.112	-0.224	0.620
NOISMN	yes	-0.253	0.598	-0.203	0.598	0.307	0.099	-0.047
ATT1MN	no	0.562	0.696	-0.542	-0.855	0.156	-0.954	0.505
CTRLMN	yes	0.291	0.352	-0.373	-0.369	0.183	-0.375	0.740
OBTMMN	no	0.496	0.023	-0.588	-0.453	0.108	-0.290	0.371
STTMMN	no	0.646	-0.331	-0.020	-0.313	0.119	-0.143	0.188
CLSSPR	yes	-0.063	0.907	-0.251	-0.001	0.408	-0.264	-0.175
ROLEPR	no	-0.231	0.859	-0.254	-0.084	0.220	-0.243	-0.192
IFLTPR	yes	0.289	0.358	-0.159	0.451	0.949	-0.199	-0.123
IFWDPR	no	-0.376	0.466	-0.064	-0.031	0.028	-0.236	-0.445
IFSTPR	yes	-0.361	-0.457	0.451	-0.169	-0.956	0.260	0.135
MAT1PR	yes	-0.438	-0.111	0.931	0.019	-0.266	0.345	-0.237
MAT2PR	yes	-0.081	-0.508	0.375	0.092	-0.244	0.822	-0.121
ATNTPR	yes	-0.041	-0.159	0.444	0.059	-0.067	-0.032	-0.031
ATINPR	yes	0.052	-0.826	0.305	0.081	-0.392	0.438	0.039
ATLRPR	no	-0.214	0.806	-0.004	-0.149	0.283	-0.360	-0.113
IFTTPR	no	-0.279	-0.102	0.532	0.432	-0.145	0.073	-0.083
IFT2PR	no	-0.422	-0.388	0.444	-0.175	-0.959	0.225	0.068
ATTTPR	no	-0.446	-0.244	0.920	-0.175	-0.398	0.319	-0.199
IFLTRP	no	0.406	0.365	-0.340	0.243	0.971	-0.221	-0.062
IFWDRP	yes	-0.191	0.449	-0.139	0.154	0.307	-0.295	-0.631
ATLRRP	no	-0.143	0.838	-0.175	-0.143	0.335	-0.408	-0.050
STTMRT	no	0.082	-0.394	0.711	0.231	-0.001	0.186	-0.247

Table 20

Reading and Mathematics Observation System - Spanish: Correlations between Scale Z-Scores and Factor Scores for Instructional Year 2 for the Bilingual Sample

Scale	Entered in Analysis	Factor Scores (N=62) QFL	DGI	ETT	NST	APC	SMT	CNT
NSTDMN	yes	-0.255	-0.197	-0.261	0.818	-0.412	0.377	0.553
CLSSMN	yes	0.561	-0.070	0.095	0.055	0.106	-0.121	-0.027
ROLEMN	yes	0.437	0.193	0.687	-0.713	0.507	0.117	0.075
SBJCMN	yes	0.489	-0.007	0.736	-0.326	0.359	0.249	0.083
IFLTMN	yes	0.576	0.505	0.094	-0.677	0.465	-0.160	-0.221
IFWDMN	no	0.083	-0.096	-0.281	-0.100	-0.203	0.153	-0.149
IFSTMN	no	0.044	0.208	0.268	-0.344	0.031	-0.103	-0.011
TECHMN	no	-0.206	0.014	-0.267	-0.034	-0.109	0.162	0.196
LANGMN	yes	-0.155	0.041	-0.290	0.111	-0.079	-0.168	-0.470
MAT1MN	yes	0.614	0.082	0.292	-0.120	0.219	0.074	-0.234
MAT2MN	yes	-0.190	-0.302	-0.097	0.483	-0.229	0.654	0.178
ATNTMN	yes	0.251	-0.271	0.756	-0.076	0.121	-0.013	0.262
ATINMN	yes	0.606	0.001	0.220	-0.291	0.045	-0.204	0.060
ATLRMN	yes	0.655	0.097	0.440	0.025	-0.030	0.150	0.186
NEN2MN	no	-0.487	0.095	-0.540	0.458	-0.284	-0.014	-0.082
PRDCMN	yes	0.294	0.355	0.231	-0.871	0.368	-0.051	-0.264
NOISMN	yes	-0.151	0.591	-0.620	0.200	-0.201	0.025	-0.165
ATT1MN	no	-	-	-	-	-	-	-
CTRLMN	yes	-0.012	-0.351	0.224	0.235	0.085	0.352	0.813
OBTMMN	no	0.415	0.296	0.119	-0.416	0.045	-0.046	-0.033
STTMMN	no	0.370	-0.058	0.260	-0.047	0.002	-0.047	0.207
CLSSPR	yes	0.348	0.801	-0.165	0.056	0.110	0.122	-0.229
ROLEPR	no	0.340	0.745	-0.132	0.100	0.064	0.173	-0.168
IFLTPR	yes	0.240	-0.010	0.521	-0.258	0.879	0.256	0.236
IFWDPR	no	0.292	0.360	0.110	-0.459	0.167	0.013	-0.442
IFSTPR	yes	-0.270	-0.341	-0.221	0.620	-0.895	-0.020	0.294
MAT1PR	yes	0.121	-0.311	0.862	-0.347	0.377	0.240	0.223
MAT2PR	yes	0.072	0.178	0.331	-0.174	0.389	0.778	0.288
ATNTPR	yes	0.132	-0.318	0.287	0.203	0.152	0.119	0.059
ATINPR	yes	0.001	-0.791	0.211	0.434	-0.132	0.365	0.699
ATLRPR	no	0.076	0.719	0.028	-0.463	0.181	-0.375	-0.634
IFTTPR	no	0.312	-0.106	0.631	-0.040	0.346	0.384	0.310
IFT2PR	no	-0.050	-0.071	-0.146	0.289	-0.820	-0.011	-0.046
ATTTPR	no	0.232	-0.360	0.756	-0.008	0.120	0.022	0.319
IFLTRP	no	0.252	0.083	0.434	-0.289	0.880	0.215	0.164
IFWDRP	yes	0.420	0.548	0.178	-0.600	0.596	0.013	-0.558
ATLRRP	no	0.065	0.777	-0.074	-0.484	0.156	-0.371	-0.653
STTMRT	no	-0.167	-0.463	0.122	0.525	-0.025	-0.012	0.250

Table 21

Reading and Mathematics Observation System - Spanish:
Correlations between Scale Z-Scores and Factor Scores
for Instructional Year 3 for the Bilingual Sample

Scale	Entered in Analysis	Factor Scores (N=20) QFL	DGI	ETT	NST	ADC	SMT	CNT
NSTDMN	yes	-0.371	0.048	-0.293	0.832	-0.407	-0.291	-0.417
CLSSMN	yes	-0.173	-0.091	0.322	0.167	-0.004	-0.091	0.308
ROLEMN	yes	0.387	0.755	0.295	-0.482	-0.401	-0.720	0.336
SBJCMN	yes	0.212	0.407	0.067	-0.156	0.467	0.303	0.554
IFLTMN	yes	0.828	0.125	0.018	-0.537	0.259	0.195	-0.146
IFWDMN	no	0.040	0.562	-0.201	-0.319	-0.232	-0.615	0.377
IFSTMN	no	0.463	0.350	0.664	-0.564	0.415	-0.159	0.113
TECHMN	no	-0.740	0.058	-0.094	0.748	-0.335	-0.393	0.148
LANGMN	yes	-0.526	0.229	-0.025	0.718	-0.373	-0.536	0.171
MAT1MN	yes	0.250	0.471	0.376	-0.427	0.155	-0.529	0.054
MAT2MN	yes	-0.235	-0.049	-0.548	0.584	0.149	0.561	0.307
ATNTMN	yes	0.050	-0.346	0.845	-0.241	-0.208	-0.207	-0.205
ATINMN	yes	0.527	-0.455	-0.295	-0.202	0.477	0.780	0.049
ATLRMN	yes	0.416	0.229	-0.317	0.040	-0.689	-0.279	-0.434
NEN2MN	no	-0.284	0.439	-0.673	0.682	-0.219	-0.048	0.108
PRDCMN	yes	0.473	-0.178	0.321	-0.811	0.125	0.130	-0.412
NOISMN	yes	0.104	0.659	-0.662	0.201	-0.016	-0.020	0.099
ATT1MN	no	-	-	-	-	-	-	-
CTRLMN	yes	-0.415	0.074	-0.173	0.470	0.275	0.134	0.774
OBTMMN	no	0.626	0.202	0.357	-0.471	-0.245	-0.532	-0.328
STTMMN	no	0.626	0.202	0.357	-0.471	-0.245	-0.532	-0.328
CLSSPR	yes	0.215	0.858	0.062	-0.058	-0.123	-0.388	-0.051
ROLEPR	no	0.203	0.845	0.065	-0.043	-0.109	-0.368	-0.074
IFLTPR	yes	-0.045	-0.036	0.216	-0.113	0.920	0.559	0.458
IFWDPR	no	0.352	-0.426	-0.301	-0.233	0.141	0.473	-0.554
IFSTPR	yes	-0.184	0.314	0.196	0.140	-0.930	-0.767	0.009
MAT1PR	yes	-0.155	-0.550	0.790	-0.123	0.079	0.000	-0.211
MAT2PR	yes	0.221	-0.593	0.203	-0.460	0.745	0.727	-0.259
ATNTPR	yes	-	-	-	-	-	-	-
ATINPR	yes	-0.053	-0.897	-0.030	0.092	0.138	0.612	-0.460
ATLRPR	no	0.075	0.852	0.208	-0.160	-0.144	-0.655	0.434
IFTTPR	no	0.028	-0.043	0.573	-0.368	0.360	0.220	0.295
IFT2PR	no	0.063	0.023	-0.010	-0.021	-0.905	-0.528	-0.403
ATTTPR	no	0.086	-0.476	0.810	-0.284	0.009	-0.011	-0.232
IFLTRP	no	-0.046	-0.070	0.108	-0.060	0.942	0.612	0.430
IFWDRP	yes	0.155	-0.550	-0.334	0.061	0.525	0.599	-0.671
ATLRRP	no	0.049	0.888	0.046	-0.090	-0.160	-0.615	0.469
STTMRT	no	-	-	-	-	-	-	-

624

Table 22

Reading and Mathematics Observation System - Spanish:
Correlations between Scale Z-Scores and Factor Scores
for Instructional Year 4 for the Bilingual Sample

		Factor Scores (N=9)						
Scale	Entered in Analysis	QFL	DGI	ETT	NST	ADC	SMT	CNT
NSTDMN	yes	-0.988	0.376	-0.979	0.838	-0.976	-0.997	-0.397
CLSSMN	yes	0.365	-0.998	0.513	0.237	0.144	0.394	0.997
ROLEMN	yes	0.804	0.191	0.709	-1.000	0.920	0.797	-0.169
SBJCMN	yes	0.278	-0.989	0.431	0.326	0.051	0.306	0.986
IFLTMN	yes	0.979	-0.542	0.999	-0.722	0.924	0.993	0.560
IFWDMN	no	-	-	-	-	-	-	-
IFSTMN	no	0.945	-0.656	0.990	-0.614	0.866	0.964	0.671
TECHMN	no	-0.781	-0.229	-0.681	0.999	-0.905	-0.773	0.207
LANGMN	yes	-0.781	-0.229	-0.681	0.999	-0.905	-0.773	0.207
MAT1MN	yes	0.325	0.722	0.173	-0.815	0.535	0.303	-0.706
MAT2MN	yes	-0.352	-0.701	-0.202	0.832	-0.559	-0.331	0.685
ATNTMN	yes	0.979	-0.542	0.999	-0.722	0.924	0.993	0.560
ATINMN	yes	-	-	-	-	-	-	-
ATLRMN	yes	-0.960	0.297	-0.954	0.875	-0.985	-0.982	-0.318
NEN2MN	no	-0.811	-0.179	-0.718	1.000	-0.925	-0.805	0.157
PRDCMN	yes	0.455	0.616	0.312	-0.889	0.649	0.436	-0.598
NOISMN	yes	-0.979	0.542	-0.999	0.722	-0.924	-0.993	-0.560
ATT1MN	no	-	-	-	-	-	-	-
CTRLMN	yes	-0.325	-0.722	-0.173	0.815	-0.535	-0.303	0.706
OBTMMN	no	0.903	-0.008	0.836	-0.980	0.977	0.902	0.030
STTMMN	no	0.903	-0.008	0.836	-0.980	0.977	0.902	0.030
CLSSPR	yes	-	-	-	-	-	-	-
ROLEPR	no	-	-	-	-	-	-	-
IFLTPR	yes	0.902	-0.745	0.966	-0.511	0.797	0.924	0.758
IFWDPR	no	-0.365	0.998	-0.513	-0.237	-0.144	-0.394	-0.997
IFSTPR	yes	-0.655	-0.394	-0.539	0.970	-0.825	-0.646	0.376
MAT1PR	yes	0.981	-0.295	0.958	-0.882	0.989	0.988	0.316
MAT2PR	yes	0.636	0.429	0.511	-0.967	0.798	0.622	-0.409
ATNTPR	yes	-	-	-	-	-	-	-
ATINPR	yes	-	-	-	-	-	-	-
ATLRPR	no	0.981	-0.295	0.958	-0.882	0.989	0.988	0.316
IFTTPR	no	0.976	-0.555	1.000	-0.711	0.918	0.991	0.573
IFT2PR	no	-0.853	0.810	-0.933	0.417	-0.731	-0.877	-0.821
ATTTPR	no	0.981	-0.295	0.958	-0.882	0.989	0.988	0.316
IFLTRP	no	0.865	-0.797	0.941	-0.438	0.746	0.888	0.808
IFWDRP	yes	-0.366	0.998	-0.513	-0.237	-0.140	-0.393	-0.997
ATLRRP	no	-	-	-	-	-	-	-
STTMRT	no	-	-	-	-	-	-	-

engaged instructional time. Both IFTTPR and ATTTPR, neither included in the factor analysis, show positive correlations with the factor in each instructional year (about .6 and .8, respectively). These are indices of the relative amounts of time devoted to literacy instruction and to instructional activity/tasks, respectively, and suggest that engaged text time increased with increases in time devoted to literacy instruction. Finally, ROLEMN and SBJCMN, both with small positive weights, show particularly low correlations in the third instructional year.

The fourth factor was named number of students (NST), although it was more complex than the corresponding factor derived in the English analysis. In the Spanish data, IFSTMN, which was not included in the factor analysis, shows a moderate negative correlation in all years (about -.5), suggesting that the factor contains a component for less explicit instruction in the meaning of connected text. A second variable not included in the analysis, NEN2MN, an index of the number of nonengaged students, is positively related to the factor (about .7) in each instructional year, which would be expected given the increased group size. Finally, OBTMMN, the length of observation, is negatively related to the factor in each year (about -.45), and suggests that larger group sizes tended to be seen in shorter observations.

The fifth factor the amount of decoding (ADC), had only two variable loadings, IFLTPR and IFSTPR, the first positive and the second negative. IFT2PR, the total percentage of time devoted to non-decoding instruction, shows a substantial negative correlation in all instructional years (about -.9), which would be expected given its high positive correlation with IFSTPR. Also, IFLTRP, which was not included in the analysis, shows a substantial positive relation in all years (about .9), again, expected given its high correlation with IFLTPR.

The sixth factor, secondary materials (SMT), had positive weights for both the quality and quantity of ancillary materials (MAT2MN and MAT2PR). ATINPR, an index of the percentage of time devoted to independent work, which did not load on this factor, nonetheless shows moderate positive correlations (about .5) in each instructional year. This suggests that the use of secondary materials is more prevalent with decreased reliance on group work (or increased reliance on seatwork).

The last factor was complex, and was labelled control (CNT) given that its strongest weighting was for the number of momentary controls. A second variable, IFWDRP, the percentage of time devoted to word meaning relative to the total percentage of time devoted to instruction on meaning (at all levels: word, sentence, and text), revealed a negative loading. In the correlation tables, IFWDPR, which was not included in the factor analysis, shows moderate negative correlations (about -.5) with the factor in all instructional years -- these would be expected given the positive correlations between IFWDRP and IFWDPR.

Correlations Between Factor Scores

Having treated the correlations between the summary values and the individual factor scores for both the English and Spanish data, next the correlations between the factor scores themselves within each instructional year are considered. Again, the correlations for English and Spanish are discussed separately, treating English first.

English Factor Score Correlations

In this section, the correlations between the English factor scores for the bilingual sample are discussed. In Table 23, these correlation coefficients are presented for each of the five instructional years.

Considering the correlation pattern found within the last four instructional years, ETT, engaged text time, shows a positive relationship (about .4) to QFL, the quality of formal language, in instructional years 2, 3, and 4. This suggests that classrooms tending toward higher amounts of engaged text time also tended toward a higher quality of formal language. Further, in the early grade levels, QFL showed a negative correlation (about -.3) with SMT, the quality and quantity of secondary materials, suggesting that the use of such ancillary materials is weakly linked to decreased formal language quality in these grade levels. The only other systematic correlations occur between DGI (direct group instruction) and PRC (productivity), which are positively correlated in each instructional year (about .4). These relationships most likely reflect the common positive correlation of ATLRRP, which was not included in the factor analysis, yet shows fairly substantial correlations with both of these factors. As such, it suggests that both direct group instruction and productivity increased with increases in the percentage of time devoted to group instruction. Other strong correlations can be found in Table 23 (especially in year 4), but these do not seem to be systematic across instructional years.

Spanish Factor Score Correlations

The correlations between the Spanish factor scores for the bilingual sample are presented in Table 24; again, these correlation coefficients are presented for each of the five instructional years, but the discussion below only treats those for the last four instructional years.

First, the correlations between factors within the last instructional year are uniformly high, which, given the high correlations between the component summary indices for this year discussed earlier, are not surprising. Consequently, the interpretation of these factors in this last year is quite difficult.

For instructional years 1 through 3, there are generally more significant relationships within the Spanish data than within the English data, and these probably reflect the generally reduced sample

Table 23

Reading and Mathematics Observation System - English:
Correlations between Factor Scores for Each Instructional Year
for the Bilingual Sample

		Factor						
		ETT	DGI	QFL	ADC	PRD	SMT	NST
	ETT	-	0.228	0.321	0.090	0.217	0.176	-0.326
	DGI		-	0.373	-0.012	0.361	0.482	-0.131
INSTRUCTIONAL	QFL			-	0.005	0.270	0.096	-0.275
YEAR 0	ADC				-	0.019	-0.197	-0.136
(N=140)	PRD					-	0.127	-0.158
	SMT						-	-0.155
	NST							-
	ETT	-	-0.028	0.082	-0.199	-0.191	-0.110	-0.210
	DGI		-	-0.060	0.471	0.553	0.169	0.075
INSTRUCTIONAL	QFL			-	-0.049	0.039	-0.251	-0.369
YEAR 1	ADC				-	0.334	0.305	0.091
(N=215)	PRD					-	0.219	-0.039
	SMT						-	0.043
	NST							-
	ETT	-	-0.130	0.363	-0.010	-0.274	-0.297	0.141
	DGI		-	0.341	-0.055	0.381	-0.157	-0.250
INSTRUCTIONAL	QFL			-	0.135	-0.119	-0.350	0.181
YEAR 2	ADC				-	-0.028	-0.265	0.014
(N=224)	PRD					-	0.160	-0.494
	SMT						-	-0.114
	NST							-
	ETT	-	0.057	0.401	0.177	0.147	-0.040	-0.181
	DGI		-	0.193	0.325	0.120	-0.104	-0.008
INSTRUCTIONAL	QFL			-	-0.060	0.174	-0.062	-0.018
YEAR 3	ADC				-	-0.326	0.102	-0.109
(n=94)	PRD					-	-0.034	0.026
	SMT						-	-0.045
	NST							-
	ETT	-	0.023	0.387	0.130	0.067	0.063	-0.135
	DGI		-	0.139	0.172	0.404	-0.514	0.458
INSTRUCTIONAL	QFL			-	0.170	-0.271	0.205	0.078
YEAR 4	ADC				-	0.100	-0.303	-0.019
(N=59)	PRD					-	-0.388	0.330
	SMT						-	-0.219
	NST							-

Table 24

Reading and Mathematics Observation System - Spanish:
Correlations between Factor Scores for Each Instructional Year
for the Bilingual Sample

		Factor						
		QFL	DGI	ETT	NST	ADC	SMT	CNT
	QFL	-	0.117	-0.224	0.031	0.185	-0.378	0.193
	DGI		-	0.418	-0.236	0.036	0.195	0.445
INSTRUCTIONAL	ETT			-	-0.258	0.008	0.494	0.518
YEAR 0	NST				-	0.309	-0.512	-0.076
(N=67)	ADC					-	-0.174	0.187
	SMT						-	0.105
	CNT							-
	QFL	-	-0.109	-0.346	-0.083	0.342	-0.326	0.292
	DGI		-	-0.179	0.078	0.429	-0.268	-0.154
INSTRUCTIONAL	ETT			-	-0.113	-0.325	0.309	-0.251
YEAR 1	NST				-	0.320	0.078	-0.345
(N=73)	ADC					-	-0.242	-0.136
	SMT						-	-0.122
	CNT							-
	QFL	-	0.244	0.389	-0.354	0.288	-0.064	-0.083
	DGI		-	-0.185	-0.306	0.192	-0.055	-0.456
INSTRUCTIONAL	ETT			-	-0.419	0.413	0.189	0.229
YEAR 2	NST				-	-0.501	0.171	0.357
(N=62)	ADC					-	0.150	-0.042
	SMT						-	0.329
	CNT							-
	QFL	-	0.220	0.152	-0.562	0.079	0.022	-0.131
	DGI		-	-0.077	-0.112	-0.194	-0.531	0.247
INSTRUCTIONAL	ETT			-	-0.411	0.003	-0.211	0.113
YEAR 3	NST				-	-0.137	0.020	-0.071
(N=20)	ADC					-	0.728	0.234
	SMT						-	-0.003
	CNT							-
	QFL	-	-0.411	0.976	-0.806	0.954	0.989	0.431
	DGI		-	-0.555	-0.188	-0.192	-0.439	-0.999
INSTRUCTIONAL	ETT			-	-0.711	0.918	0.990	0.574
YEAR 4	NST				-	-0.921	-0.799	0.167
(N=9)	ADC					-	0.961	0.212
	SMT						-	0.458
	CNT							-

size. Of those relationships, only two appear to be systematic across instructional years. Both the quality of formal language (QFL) and engaged text time (ETT) are negatively related to the number of students (NST). This indicates that in the Spanish classrooms, the quality of formal language increased with decreases in group size, as did engaged text time.

This concludes the discussion of the RAMOS data collected in English and Spanish reading classrooms. As was seen, the structure of the two data sets was quite similar, which perhaps would be expected given that many of the teachers observed taught both English and Spanish reading, albeit to different students. In the next sections, data collected from the second source of information on reading instruction, the Teacher Checklists, are discussed.

TEACHER CHECKLISTS

A secondary source of instructional data was obtained from regular interviews with the teachers who were primarily responsible for providing reading instruction to each of the target students. In this section, as in that treating the RAMOS, the structure of the interview instrument is first described, followed by discussions of the data collection and processing procedures, the derivation of summary indices, and descriptive statistics on the aggregated measures for the bilingual sample. These are followed by descriptions of further data reduction procedures which yielded factor scores, and the descriptive statistics and inter-correlations associated with them.

Instrument Description

To supplement and verify the information obtained in the RAMOS observations, the teachers in the study were interviewed bimonthly (on average) during the school year by SEDL research staff using the Teacher Checklist. Employing the same codes used in the RAMOS, the Checklist essentially obtains from each of the teachers their plans regarding reading instruction for a two-week period for each of the target students taught. For each student, the strategies or skills that were to be taught during the two-week period were listed; and for each strategy or skill listed, the following were indicated: instructional focus, material (type, title, and section of the book), type of activity, language of instruction, instructor (e.g., teacher or teacher aide), role of the instructor, and total minutes devoted to the teaching of the strategy/skill over the two-week period.

In Appendix G, the Teacher Instructional Plan (the Checklist coding sheet) and the Checklist Master Code Sheet are presented. Concerning the former, the top of the page is used for recording basic identification information: the teacher interviewed, the associated school and grade level, the dates covered by the instructional plans described, and the identification and group assignments of the target students. The specific instructional categories appearing in the bottom half of the page for delineating specific instructional plans are summarized below (refer to the Checklist Master Code Sheet in Appendix G for the specific codes allowed under each category):

Group/Students: the unique identification number of the group (or specific target student if instruction is individualized) whose instructional plans are being described.

Strategy Number: the nominal number of the particular instructional strategy being described for the group under consideration.

Instructional Focus: the instructional emphasis associated with the particular strategy being described (under six skill headings of general, grammar, vocabulary, decoding, comprehension, and interpretation skills).

Materials: the materials being used in the delivery of the particular strategy described (allowing two distinct codes for each strategy), providing the following information:

Type: the type of material used;

Title: if a text is being used, its title;

Page(s): if a text is being used, the page numbers covered.

Activity: the type of activity used as a vehicle to convey the instructional content of the strategy described (under eight activity headings of lecture, discussion, independent work, questions/answers, recitation, audio-visual, other reading-related activities, and non-reading activities.

Language: the language employed by the instructor in the delivery of the instructional strategy described.

Instructor: the classification of the instructor with primary responsibility for overseeing the delivery of the instructional strategy being described (e.g., teacher, aide, volunteer).

Role: the role of the instructor in delivering the instructional strategy being described.

Total Minutes/Two-week Period: the total number of minutes in the two-week period being considered which is to be devoted to the particular instructional strategy described.

The Checklist and RAMOS codings are identical for those overlapping categories that are covered, with one exception: given the non-instructional nature of transitional activities, their coding, which was allowed in the RAMOS (under the Activity/Task category), was not possible in the Checklist coding (under the analogous Activity category).

An example taken from a completed protocol will demonstrate the method this system employs in characterizing instructional plans. Figure 8 displays a completed Checklist coding sheet taken from a Site 5, second-grade classroom during the last year of data collection. It is an interview with the same teacher whose RAMOS coding sheet was used as an example in the previous section, and covers the same time period. This example will be the basis of further examples used in describing the derivation of instructional summary indices, so a considerable amount of detail will be given here.

At the top of the form, this teacher's normal groupings for reading instruction are given: Group 1 contains five students, two of whom are target students in this study (with unique identification numbers subsequently provided); Group 2 contains four students, where one is a target; and Group 3 contains six students, one being a target. In the lower portion of the coding sheet, the three groups are listed along with a specification of the major instructional strategies used by the teacher over the two-week course covered by the interview. Looking at the materials columns first, note that two groups are using the Magic Times readers (from the Macmillan series): Group 1 is working in the last sections of the book, while Group 2 is working in the middle sections ('E' and 'M' under the column labelled Page(s), respectively). The third group is in the beginning sections of Rainbow World (indicated by a 'B'). Thus, from the difficulty sequencing of this material, Group 1 constitutes this classroom's top reading group, Group 2 the middle group, and Group 3 the bottom group.

Before considering the individual instruction planned for the three groups, note that English is the only language of instruction ('E' under Language). Further, the teacher is always given as the instructor primarily responsible for the instruction delivered ('T' under Instructor). Thus, this classroom offers only English reading, and employs no supporting instructors (e.g., resource teachers, teacher aides), at least not in primary roles.

Considering the instruction offered Group 1, first note that six strategies are listed (the six entries under the column labelled Strategy Number). The first strategy for Group 1 is focused on the comprehension of literal facts from text ('CLF' under Instructional Focus), and relies on the basal reader ('BR' under Materials: Type). The instructional activity is reading and responding ('RRR' under Activity), where the instructor's role is one of direct instruction ('N' under Role); this strategy occupies approximately 50 minutes of instructional time over the two-week period of interest (about five minutes per day). The second strategy is focused on whole text recognition ('DWT'), and also employs the basal reader ('BR'). The activity is reading aloud ('RRA'), and as before, is conducted by the teacher in a direct instruction mode; it occupies about 15 minutes per day. The third strategy's focus is on letter cluster-sound recognition ('DCS'), and employes a supplementary book ('SB'). The

TEACHER INSTRUCTIONAL PLAN

Teacher [redacted]

School [redacted]

Grade 2

Period 1/20/83 to 2/3/83
(date) (date)

READING GROUPS:

Group Name or I.D. No.	No. of Students	INITIALS OF STUDENTS IN GROUP 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	5	GG 5[illegible]	RG 5[illegible]																		
2	4	AE 5[illegible]																			
3	6	AS 5259																			

GROUP/ STUDENTS	Strategy Number	Instructional Focus	MATERIALS Type	Title	Page(s)	Activity	Language	Instructor	Role	Total Minutes/ Two-week Period	SPECIAL NOTES
1	1	CLF	BR	MAGIC TIMES	E	RRR	E	T	N	50	
	2	DWT	BR			RRA	E	T	N	150	
	3	DCS	SB			QSA	E	T	N	50	
	4	DCS	BW			WWA	E	T	F	100	
	5	VVE	SH			QSA	E	T	N	50	
	6	VVE	PP			WCR	E	T	F	200	
2	1	CLF	BR	MAGIC TIMES	M	RRR	E	T	N	50	
	2	DWT	BR			RRA	E	T	N	150	
	3	DCS	SB			QSA	E	T	F	50	
	4	DCS	BW			WWA	E	T	N	100	
	5	VVE	SH			QSA	E	T	N	60	
	6	VVE	PP			WCR	E	T	F	200	
3	1	DSR	SC	RAINBOW WORLD	B	RRA	E	T	N	30	
	2	DWW	BR			QSA	E	T	N	30	
	3	DWT	BR			RRA	E	T	N	200	
	4	CLF	BR			RRR	E	T	N	80	
	5	DCS	SB			QSA	E	T	N	50	
	6	DCS	BW			WWA	E	T	F	100	

Figure 8. Checklist: Sample raw protocol from a Site 5 second grade.

activity consists of questions with spoken answers ('QSA'); it is conducted by the teacher in a direct instruction role, and occupies about five minutes daily. The fourth activity also focuses on letter cluster-sound recognition, but employs a basal workbook ('BW'), where the activity consists of independent work with written answers ('WWA'). Again, the teacher has primary responsibility for overseeing the work, playing the role of facilitator. The activity occupies about 10 minutes per day. The final two strategies both are concerned with vocabulary enrichment (in English). The first (lasting about five minutes per day) consists of group work with spoken answers to questions from the teacher in a direct instruction role. The second consists of 20 minutes per day of independent creative writing, supervised by the teacher in a role of facilitation. Thus, overall the first group's instructional plans call for about 60 minutes of reading instruction daily, 20 minutes devoted to whole text work, 15 minutes to decoding work, and 25 minutes to vocabulary work.

The instruction planned for Group 2, the middle reading group, is very similar to that planned for Group 1. Indeed, the only differences concern (1) their place within the basal series employed (discussed earlier), and (2) the teacher's role during instruction on letter cluster-sound recognition.

Group 3, the bottom reading group, has about 10 minutes less planned instruction per day in reading. This group receives similar instruction in both decoding and whole text work, though the latter is extended by about half again as much time. This group does not receive any of the vocabulary work given to the other two groups, but rather a small amount of instruction devoted to whole word ('DWW') and sentence ('DSR') recognition.

Thus, the information obtained from the Checklist, though not as detailed as that obtained from the RAMOS, nonetheless provides specific information in a RAMOS-like format about the planned instruction for individual reading groups. Before discussing the derivation of summary indices from these raw protocols, the data collection and processing procedures will be discussed.

Data Collection and Processing

In the initial year of the study, Checklists were completed monthly by trained field observers at the single site (Site 0) then participating in the study. In all subsequent years, these interviews were conducted by trained in-house SEDL staff during their regular site visits (generally, in November/December, January/February, and April/May). As the interviews were completed and returned to the Laboratory, they were checked by in-house staff for coding completeness and accuracy. Once checked, the raw protocols were then entered into computer files, frequency analyses conducted, and identified coding/entry errors corrected. As detailed below, the raw alphanumeric codes were then transformed to ordinal numeric codes based on scaling routines, and summary indices of the instructional dimensions of interest for each target student represented in each Checklist interview were generated. These files were then segregated into English and Spanish reading summaries. At the end of the data collection phase of the study, these language-specific collection year summary files were then merged to create language-specific instructional year summary files (i.e., English and Spanish summary files organized by instructional year, kindergarten through fourth grade).

Segregation of English and Spanish Protocols

The segregation of individual student Checklist summaries into those representing English reading and Spanish reading was made, as in the RAMOS segregation procedure, mainly on the basis of the primary language of the materials used in the instructional groups each target student was assigned. Similar information concerning the language of the materials as indicated on any RAMOS protocols obtained during the same period of time were also reviewed, and any discrepancies between these sources of information were checked with the field observer (or relevant teacher) to determine whether a given student at a given time was primarily enrolled in an English, Spanish, or dual language reading program. We now turn to a discussion of the procedures used to summarize the information contained on the English and Spanish reading Checklists.

Derivation of Summary Measures

As mentioned earlier, the first step in summarizing the information contained on the raw Checklist interview forms was to enter them into computer files. Figure 9 displays the straightforward computer coding for the sample protocol appearing in Figure 8. The first four columns give identification information, uniquely specifying the site, school, teacher, and record. The remaining columns delineate the instructional strategy, one line for each strategy encountered under each group, specifying the strategy number, the instructional focus, the primary material (both its type, and if a basal text, a code uniquely identifying the particular text and the group's relative position in it), the activity, the language of instruction, the classification and role of the instructor, and the number of minutes devoted to the strategy. Finally, with each interview coding, a table (not depicted in Figure 9) is provided which specifies the reading group number associated with each target student discussed in the interview.

Instructional dimensions were quantified by constructing scales from the raw codes under each of the Checklist categories. The scales derived were a subset of those employed in the RAMOS scaling; the values assigned individual codes under each of these scales were discussed earlier (see the RAMOS Derivation of Summary Measures, and the Appendix C Definition of Scales). The instructional dimensions derived for the Checklists are summarized below:

							MATERIAL:						
SITE	SCHOOL	TEACHER	RECORD	GROUP	STRATEGY	INST FOCUS	TYPE	TITLE	ACTIVITY	LANGUAGE	CLASSFCTN	ROLE	MINUTES
5	20	202	201	01	01	CLF	BR	10606E	RRR	E	T	M	050
5	20	202	202	01	02	DWT	BR		RRA	E	T	M	150
5	20	202	203	01	03	DCS	SB		QSA	E	T	M	050
5	20	202	204	01	04	DCS	BW		WWA	E	T	F	100
5	20	202	205	01	05	VVE	SN		QSA	E	T	M	050
5	20	202	206	01	06	VVE	PP		WCR	E	T	F	200
5	20	202	207	02	01	CLF	BR	10606M	RRR	E	T	M	050
5	20	202	208	02	02	DWT	BR		RRA	E	T	M	150
5	20	202	209	02	03	DCS	SB		QSA	E	T	F	050
5	20	202	210	02	04	DCS	BW		WWA	E	T	M	100
5	20	202	211	02	05	VVE	[illegible]		QSA	E	T	M	050
5	20	202	212	02	06	VVE	PP		WCR	E	T	F	200
5	20	202	213	03	01	DSR	SC		RRA	E	T	M	030
5	20	202	214	03	02	DWW	BR	10605B	QSA	E	T	M	030
5	20	202	215	03	03	DWT	BR		RPA	E	T	M	200
5	20	202	216	03	04	CLF	BR		RRR	E	T	M	080
5	20	202	217	03	[illegible]	DCS	SB		QSA	E	T	M	050
5	20	202	218	03	06	DCS	BW		WWA	E	T	F	100

Figure 9. Checklist coding based on Site 5 second grade sample protocol.

637

Instructional Focus: the relative explicitness of the instructional strategies employed in three instructional subcategories:

Letter-Sound Unit: the relative explicitness of the instructional emphasis placed on decoding, ranging from work on isolated units (auditory discrimination, letter recognition, letter-name work) to non-explicit letter-sound pairing (whole word recognition, spelling practice) to explicit letter-sound pairing (letter cluster-sound recognition, letter-sound recognition, spelling pattern recognition).

Word Unit - Meaning: the relative explicitness of the instructional emphasis placed on word meaning, ranging from low (dictionary usage) to mid-level (noun derivative, compound words) to high (antonyms/synonyms, vocabulary enrichment).

Sentence and Text Units - Meaning: the relative explicitness of the instructional emphasis placed on sentence and text meaning, ranging from low (literal facts) to mid-level (story sequence, predicting events) to high (major ideas, making inferences).

Materials (Primary and Ancillary): the amount of text contained in the materials used, ranging from minimal (art material, tape recorder) to mid-level (phrase card, chalkboard) to substantial (basal reader, library book). [Note: Ancillary materials were rarely mentioned in the interviews, and were not subsequently analyzed.]

Number of Basals (not included in the RAMOS scales): the number of different basals used in the delivery of the instruction.

Activity/Task: the level of formal language demand required by particular activity/tasks in two instructional subcategories:

Independent: the level of formal language demand for activity/tasks classified as independent work, ranging from minimal (art activity, copying material) to mid-level (writing from dictation, writing answers) to substantial (test taking, creating writing).

Listening and Responding in Group: the level of formal language demand for activity/tasks classified as listening and responding in groups, ranging from minimal (music activity, playing games) to mid-level (watch-listen, listen-story) to substantial (listen-lecture, discussion-speak).

Classification: the level of the instructor's formal training, ranging from minimal (volunteer) to mid-level (teacher aide) to substantial (substitute teacher, resource teacher, teacher).

Role: the level of formal instruction provided, ranging from minimal (preparation, control, management) to mid-level (facilitation) to substantial (direct instruction).

Number of Students: the number of students contained in the instructional group.

Time: the total amount of time devoted to the described strategies for the group under consideration.

Rank (not included in the RAMOS scales): the relative position of the target students' reading group with respect to the following criteria:

> Internal: the relative ranking of the target students' reading group with respect to the other reading groups of the classroom, ranging from low (one of the lowest reading groups) to mid-level (the average reading group) to high (one of the top reading groups).
>
> External: the relative ranking of the target student's reading group with respect to the grade level expectations of the basal reading series employed, ranging from low (below grade level expectations) to mid-level (at grade level expectations) to high (above grade level expectations).

In summarizing a given student's planned instruction for a particular interview, a set of summary indices based on the above scales were computed. First, for each scale the mean value was computed as an index of the quality of the instructional dimension. This was done by simply converting the alphanumeric coding for a particular value to its numeric scale equivalent, multiplying by the number of applicable minutes, then summing all such transformed values for the particular scale and dividing by the total number of applicable minutes. For example, for Group 1 in Figure 9, 'DWT' in the second line under the heading INST FOCUS has a numeric value of '2' under the scale Instructional Focus: Letter Sound Units, and contributes 150 minutes of this value to the total scale sum. The other remaining codes which apply to this scale for this group appear in the third and fourtn lines -- 'DCS' has a numeric value of '3', and contributes a total of 50 and 100 minutes, respectively. Thus, the mean value is:

$$[(2 * 150) + (3 * 50) + (3 * 100)] / (150 + 50 + 100) = 2.5$$

As an index of the variability of the quality of instruction under the instructional dimensions of interest, the standard deviation of the scaled values around the mean value was computed. Finally, the percent of time in which applicable codes were found, relative to the total amount of planned instructional time for the student, was computed as an index of the quantity of the instructional dimension for the interview period covered.

As an illustration, the sets of summary indices for the four target students identified in the sample protocol are displayed in Figure 10. The first five columns give student-protocol identification information (student identification number under STUDENT, teacher identification number under TEACHER, group number under GROUP, record number under RECORD, and interview date under DATE). The remaining columns to the right delineate the scale measures, presented under mnemonic headings. The summary indices computed for target student 5099, a member of the top reading group, are contained on the first three lines of Figure 10. The first set of entries for this student are under the category of Instructional Focus: Letter-Sound Unit (LETTR-SND), and these values are 2.5, 0.5, and 50.0, representing the mean (record number 1), standard deviation (record number 2), and percent of time (record number 3) which summarize the planned instructional focus for decoding for this student. Thus, the average quality of the planned decoding instruction was 2.5, indicating fairly explicit instruction in decoding (see the formula above). The standard deviation of 0.5 indicates that there was some degree of fluctuation in the quality of decoding instruction. The percent value of 50.0 indicates that planned decoding instruction accounted for half the instructional focus time.

The interested reader is left to track the other summary indices, but a few points of particular interest should be made. First, as in the RAMOS summary indices, the value '99.9' was used in the percentage entries instead of '100.0' in an effort to conserve file space (but again, such values were converted to 100.0 in all subsequent analyses). Second, values of '-1.0' indicate non-applicable values (e.g., neither standard deviations nor percentages of time are computed for the total number of basals indicated). Third, the coding of the internal and external rankings are given simply as 'L', 'M', and 'H', for low, medium, and high, respectively.

One final point of interest concerns the uniqueness of information with respect to individual students. In the observation data, identical sets of summary indices were rare, since (1) student groupings generally fluctuated within a given observation, and (2) within a given group, information was usually obtained about how the activities of individual students differed from each other. However, for the Checklist data, identical sets will be more common since most teachers gave their instructional plans based on groups of students rather than individuals. Thus, note that in Figure 10, the summary information for target students 5109 and 5120 are identical, as both are members of the second reading group.

The summary indices just described are the basic measures of planned instruction employed in this study, yet they are not the end-product used in the analyses integrating reading and instruction. The derivation of these measures is the subject of the next section.

STUDENT	TEACHER	GROUP	RECORD	DATE	INSTRUCTIONAL FOCUS: LETTR-SND	WORD	SENT/TEXT	MATERIAL	# OF BASALS	ACTIVITY-TASK: INDPNDNT	LISTN-RESP	CLSSFCTN	ROLE	# OF STUDENTS	TIME	RANK: INTERNAL	EXTERNAL
5099	202	2	1	012083	2.5	3.0	1.0	6.3	1.0	2.7	2.5	7.0	7.7	4	600	H	L
5099	202	2	2	012083	0.5	0.0	0.0	1.8	-1.0	0.5	0.5	0.0	1.5				
5099	202	2	3	012083	50.0	41 7	8.3	99.9	-1.0	50.0	50.0	99.9	99.5				
5109	202	1	1	012083	2.5	3.0	1.0	6.3	1.0	7.7	2.5	7.0	7.5	5	600	H	M
5109	202	1	2	012083	0.5	0.0	0.0	1.8	-1.0	0.5	0.5	0.0	1.5				
5109	202	1	3	012083	50.0	41.7	8.3	99.9	-1.0	50.0	50.0	99.9	99.5				
5120	202	1	1	012083	2.5	3.0	1.0	6.3	1.0	2.7	2.5	7.0	7.5	5	600	H	M
5120	202	1	2	012083	0.5	0.0	0.0	1.8	-1.0	0.5	0.5	0.0	1.5				
5120	202	1	3	012083	50.0	41.7	8.3	99.9	-1.0	50.0	50.0	99 9	99.5				
5259	202	3	1	012083	2.4	-1.0	1.0	7.6	1.0	2.0	2.4	7.0	8.4	6	490	L	L
5259	202	3	2	012083	0.5	-1.0	0.0	1.0	-1.0	0.0	0.5	0.0	1.2				
5259	202	3	3	012083	83.7	-1.0	16.3	99.9	-1.0	20.4	79.6	99.9	99.9				

82

Figure 10. Checklist derived summary indices for four target students from the Site 5 second grade sample protocol.

Aggregation Procedures

As in the aggregation of the RAMOS individual summaries, the average scale values obtained from individual Checklist summaries (as just described) were next themselves averaged within semester for each individual target student (i.e., all protocol averages for a given student within a given semester were collected, then averaged). Finally, the Fall and Spring average semester values were averaged to obtain yearly average values for each target student for each scale (thus giving equal weighting to Fall and Spring averaged summaries in cases where the numbers of Checklists, or the amounts of time represented within them, for a given student, were not equally distributed across the two semesters). The interpretations of the resulting quality and quantity values remain unchanged: for a given scale, the quality index represents the average scale value and the quantity index represents the average percent of student time planned in the instructional dimension relevant to that scale. These values are the basis of most of the subsequent analyses.

Descriptive Statistics

In this section, the descriptive statistics on the Checklist summary indices for the bilingual sample are described, treating English first, then Spanish.

Instruction in English Reading

For the English reading summaries, Tables 25 and 26 display the descriptive statistics (mean, standard deviation, and the number of cases) for the average quality and quantity indices of instruction, respectively. The data are summarized for each of the five instructional years, and the instructional scales appear under mnemonic headings. For the quality indices (Table 25), the following four-character scale names, identical for those scales shared with the set of RAMOS scales, are used (the characters MN appended to these names stand for Mean):

IFLT: Instructional Focus: Letter-Sound Unit
IFWD: Instructional Focus: Word Unit - Meaning
IFST: Instructional Focus: Sentence/Text Unit - Meaning
MAT1: Primary Materials
NBRD: Number of Basals
ATIN: Activity/Task: Independent
ATLR: Activity/Task: Listening and Responding in Group
CLSS: Instructor: Classification
ROLE: Instructor: Role
NSTD: Number of Students
TIME: Total Strategy Time
RNKI: Rank: Internal
RNKE: Rank: External

In addition to the above values, the number of Checklists averaged within semester (NCISMN) is also tabled.

Table 25

Checklists - English:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for the Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTMN	M	1.7	2.2	2.2	2.2	2.0
IFLTMN	S	0.6	0.3	0.2	0.2	0.1
IFLTMN	N	161	193	228	97	60
IFWDMN	M	3.0	2.9	2.5	2.4	2.4
IFWDMN	S	0.1	0.2	0.5	0.6	0.7
IFWDMN	N	130	89	171	88	59
IFSTMN	M	2.3	2.0	2.0	2.1	1.8
IFSTMN	S	0.7	0.6	0.6	0.5	0.5
IFSTMN	N	142	184	211	93	60
MAT1MN	M	4.4	5.8	7.0	7.4	7.5
MAT1MN	S	1.4	1.4	0.6	0.4	0.3
MAT1MN	N	163	203	228	97	60
NBRDMN	M	0.7	0.9	1.1	1.1	1.0
NBRDMN	S	0.6	0.5	0.4	0.4	0.1
NBRDMN	N	163	203	228	97	60
ATINMN	M	1.6	2.0	2.0	2.2	2.1
ATINMN	S	0.7	0.4	0.2	0.3	0.2
ATINMN	N	144	187	222	94	60
ATLRMN	M	2.3	2.3	2.3	2.5	2.4
ATLRMN	S	0.4	0.3	0.3	0.2	0.2
ATLRMN	N	163	195	228	97	[illegible]
CLSSMN	M	6.6	6.9	7.0	6.9	6.8
CLSSMN	S	0.5	0.2	0.1	0.1	0.6
CLSSMN	N	163	203	228	97	60
ROLEMN	M	8.2	8.1	7.8	7.9	7.7
ROLEMN	S	0.5	0.7	0.6	0.5	0.5
ROLEMN	N	163	203	228	97	60
NSTDMN	M	10.3	8.3	9.0	11.7	14.1
NSTLMN	S	5.9	2.4	3.3	4.5	5.6
NSTDMN	N	163	203	228	97	60
TIMEMN	M	370.3	396.2	398.0	636.2	586.6
TIMEMN	S	145.2	181.6	232.4	175.1	172.7
TIMEMN	N	163	203	228	97	60
RNKIMN	M	2.7	2.0	2.0	2.2	2.1
RNKIMN	S	0.5	0.8	0.8	0.8	0.9
RNKIMN	N	16	106	179	70	30
RNKEMN	M	2.1	1.4	1.5	1.7	1.7
RNKEMN	S	0.5	0.5	0.5	0.6	0.4
RNKEMN	N	65	180	228	97	60
NCISMN	M	2.2	2.6	1.7	1.4	1.5
NCISMN	S	2.0	2.3	0.9	0.2	0.2
NCISMN	N	163	203	228	97	60

Table 26

Checklists - English:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for the Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTPR	M	57.6	75.5	70.1	57.1	44.4
IFLTPR	S	24.7	21.8	13.5	15.8	15.7
IFLTPR	N	163	203	228	97	60
IFWDPR	M	22.9	5.0	9.5	15.1	19.7
IFWDPR	S	19.8	9.4	9.3	9.6	13.6
IFWDPR	N	163	203	228	97	60
IFSTPR	M	19.6	22.4	20.4	27.8	35.8
IFSTPR	S	16.9	21.4	11.4	15.4	17.9
IFSTPR	N	163	203	228	97	60
ATINPR	M	21.5	30.8	30.0	29.0	32.3
ATINPR	S	14.1	23.7	14.5	14.2	10.9
ATINPR	N	163	203	228	97	60
ATLRPR	M	78.4	69.1	70.0	71.0	67.6
ATLRPR	S	14.1	23.7	14.5	14.2	10.9
ATLRPR	N	163	203	228	97	60
IFTZPR	M	42.4	27.4	29.9	42.9	55.6
IFTZPR	S	24.7	21.8	13.5	15.8	15.7
IFTZPR	N	163	203	228	97	60
IFWDRP	M	51.4	17.6	30.2	35.5	34.2
IFWDRP	S	30.9	25.3	26.0	22.1	20.0
IFWDRP	N	154	192	217	95	60

The quantity indices of English reading instruction are presented in Table 26. Given that the Checklist format required a complete listing across each of the categories for each strategy given by the interviewed teacher, the percentage of time associated with many of the categories is either (1) constant at 100% (for the variables of Materials, Classification, Role, and Number of Students), or (2) inappropriate (for the variables of Number of Basals, Time, Rank, and Number of Checklists in Semester). Accordingly, only those percents of time associated with the multiple categories under Instructional Focus (IFLT, IFWD, and IFST) and Activity/Task (ATIN and ATLR) are tabled. Two additional percentages derived from the individual Instructional Focus percentages are also tabled. First, IFT2PR represents the average percent of literacy-focused instructional time which is not devoted to decoding instruction (i.e., the averaged total of the percent of time associated with the instructional focus categories of IFWDPR and IFSTPR). Second, the relative percentage appearing as IFWDRP is the percentage of time devoted to word meanings (IFWDPR) relative to the percentage of non-decoding time (IFT2PR -- the percentage of time spend on word, sentence, and text meaning).

In describing the English reading instruction as represented in these tables, only instructional years 1 through 4 will be considered, since these data are the main focus of subsequent analyses. Note that the number of observations at each year reflects the cohort structure of the study (see Volume 2), and thus, the number of data points at instructional years 3 and 4 (and also at year 0) is reduced from that available at years 1 and 2, requiring caution in their interpretation. Such interpretation is further complicated by the loss of students in Spanish reading programs: as students gain fluency in English oral language skills and Spanish reading skills, they exit the Spanish reading programs for assignment to exclusive English reading programs.

First, considering the instructional focus categories, about three-fourths of the instructional time planned was devoted to decoding instruction during the first and second instructional years, falling by about 15% in each of the subsequent two years. This instruction was to be focused on non-explicit letter-sound pairings in each of the grades -- the relatively low variability suggests that little explicit letter-sound work was planned. The amount of time to be spent on word meanings was around 5% in first grade, increasing by about 5% with each subsequent year. The instruction that was planned was quite explicit in the first year, with decreasing explicitness in each of the subsequent years (accompanied by increasing variability). About 20% of the planned instructional focus was devoted to instruction in the meanings of sentences and texts for the first two years, increasing by about 8% in each of the two subsequent years -- the large variability found during the first and last year, however, suggests substantial individual differences in the amounts of time devoted to this instructional focus. The quality of the planned instruction was fairly stable across the four years (though somewhat lower in the third year), and was generally non-explicit (e.g., favoring a focus on literal facts over making inferences).

For materials, there is a clear trend toward more planned text usage in years 2-4 as compared to years 0-1, though the variability is twice the value in these early years as compared to the later years. The number of basals employed is fairly stable across the years, with an average value of one.

For the activity/task scales, about 30% of the planned instructional activities consist of independent work, and this value is fairly stable across instructional years 1-4 (though again, there is substantial variability associated with the first instructional year index). Conversely, group work represented about 70% of the planned activities. For independent work, the level of formal language demand was, on average, mid-level and fairly stable across the years. Similarly, for group work, the level of formal language demand was also relatively stable across years, however, it is noticeably higher.

For the instructor scales, the instructor primarily responsible for the delivery of the planned instruction was the teacher in each of the four instructional years. Further, the role played was also stable across the four instructional years, and tended to be mid-way between facilitation and direct instruction.

Concerning the number of students expected in each group, there appears to be a trend toward larger groups, with an average of about nine students in the first years, and about 13 in the later years.

For the time measures, the average amount of planned daily reading instruction is about 40 minutes in the initial year, increasing to about 60 minutes in the later years. Note, however, the relatively large variance in the second year.

For the reading group rankings, the internal rank is about average across the four years, indicating that the target student sample taken in this study is representative of the average reading groups contained within each of the classrooms. The external ranking is below average in each year, which matches the English reading comprehension data discussed in Volume 5.

Finally, the number of Checklists averaged within a semester shows the trend expected from the change in data collection schedule discussed earlier: on average, three interviews were conducted with each teacher in each year.

The English quality and quantity indices for each individual site are presented in Appendix H, which contains 10 tables, two for each of the five sites (the first of a pair presenting the quality indices, and the second, the quantity indices). As in the RAMOS data, some site differences are apparent, but these will not be discussed here.

Instruction in Spanish Reading

For the Spanish reading summaries, Tables 27 and 28 display the descriptive statistics for the average quality and quantity indices of

Table 27

Checklists - Spanish:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for the Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTMN	M	2.2	2.4	2.3	2.2	2.1
IFLTMN	S	0.6	0.3	0.2	0.3	0.1
IFLTMN	N	87	75	90	48	20
IFWDMN	M	3.0	2.8	2.7	2.8	3.0
IFWDMN	S	0.0	0.6	0.5	0.4	0.0
IFWDMN	N	71	24	38	35	10
IFSTMN	M	1.8	1.8	1.8	1.2	1.8
IFSTMN	S	0.4	0.7	0.7	0.4	0.9
IFSTMN	N	59	56	74	32	20
MAT1MN	M	4.2	6.0	7.0	7.6	7.2
MAT1MN	S	1.4	1.1	0.7	0.5	1.4
MAT1MN	N	93	75	90	48	21
NBRDMN	M	0.7	0.9	1.0	0.8	0.5
NBRDMN	S	0.4	0.2	0.1	0.5	0.5
NBRDMN	N	93	75	90	48	21
ATINMN	M	2.0	1.9	2.0	2.2	2.4
ATINMN	S	0.9	0.6	0.2	0.2	0.5
ATINMN	N	57	68	80	23	20
ATLRMN	M	2.1	2.2	2.4	2.3	2.4
ATLRMN	S	0.4	0.4	0.3	0.4	0.3
ATLRMN	N	93	75	90	48	21
CLSSMN	M	6.8	6.9	6.9	7.0	6.8
CLSSMN	S	0.4	0.2	0.2	0.2	0.4
CLSSMN	N	93	75	90	48	21
ROLEMN	M	8.4	8.3	8.0	8.6	8.7
ROLEMN	S	0.6	0.5	0.5	0.5	0.3
ROLEMN	N	93	75	90	48	21
NSTDMN	M	7.4	10.2	8.9	11.9	13.6
NSTDMN	S	2.8	5.1	4.3	5.7	7.4
NSTDMN	N	93	75	90	48	21
TIMEMN	M	226.7	317.9	402.7	299.3	341.2
TIMEMN	S	98.7	154.1	234.6	186.3	94.5
TIMEMN	N	93	75	90	48	21
RNKIMN	M	1.9	2.0	2.1	1.8	1.9
RNKIMN	S	0.8	0.8	0.5	0.4	0.2
RNKIMN	N	8	28	64	34	18
RMKEMN	M	1.7	1.5	1.6	1.3	1.3
RMKEMN	S	0.5	0.4	0.6	0.5	0.3
RMKEMN	N	46	61	76	47	20
NCISMN	M	2.9	2.5	1.7	1.4	1.3
NCISMN	S	2.8	[illegible]	0.8	0.5	0.3
NCISMN	N	93	75	90	48	21

648

Table 28

Checklists - Spanish:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for the Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTPR	M	61.2	81.8	77.9	68.0	60.2
IFLTPR	S	27.6	15.7	14.3	20.2	20.0
IFLTPR	N	93	75	90	48	21
IFWDPR	M	26.2	4.5	5.8	16.8	14.0
IFWDPR	S	24.0	8.2	10.4	15.6	15.4
IFWDPR	N	93	75	96	48	21
IFSTPR	M	12.5	13.6	16.2	15.2	25.7
IFSTPR	S	14.5	12.5	12.8	16.2	18.1
IFSTPR	N	93	75	90	48	21
ATINPR	M	14.0	22.5	28.2	11.2	16.8
ATINPR	S	14.3	15.1	14.8	14.0	6.6
ATINPR	N	93	75	90	48	21
ATLRPR	M	85.9	77.5	71.8	89.8	83.1
ATLRPR	S	14.3	15.1	14.8	14.0	6.6
ATLRPR	N	93	75	90	48	21
IFT2PR	M	38.8	18.1	22.1	32.0	39.7
IFT2PR	S	27.6	15.7	14.3	20.2	20.0
IFT2PR	N	93	75	90	48	21
IFWDRP	M	57.5	19.1	25.9	49.0	29.5
IFWDRP	S	35.8	29.3	34.2	40.0	33.0
IFWDRP	N	83	58	81	46	21

instruction, respectively. These data are also summarized for each of the five instructional years, and are structurally identical to those displaying the English data. As in the above discussion, the kindergarten data will not be discussed here; and again, the differences dues to cohort require cautious interpretation.

Considering the instructional focus categories, about three-quarters of the instructional time planned was devoted to decoding instruction during the first and second years, matching the planned English instruction. However, whereas this value fell to about 45% in the English data, it dropped only to about 60% in the Spanish data. The explicitness of this Spanish decoding instruction matched that of the English instruction: a focus on non-explicit letter-sound pairings in each of the grades, with little explicit letter-sound work planned. The amount of time to be spent on word meanings was around 5% in first and second grade, and around 15% in third and fourth grade, thus, showing a different pattern than that found in English. The instruction planned was quite explicit in the first grade, and did not show the steady decline in explicitness found in the English data. About 15% of the planned instructional focus was devoted to instruction in the meanings of sentences and texts for the first three years, increasing by about 10% in the fourth year -- for these students, the relatively larger amounts of time devoted to decoding instruction were associated with relatively smaller amounts of time devoted to the study of connected text. The explicitness of this instruction is fairly constant across instructional year , and tends to be somewhat lower than that found in English.

For materials, as in English, there is a clear trend toward more planned text usage in years 2-4 as compared to years 0-1, though again, the variability in the early years is substantial relative to the later years. Also, the number of basals employed has an average value of one, though the usage of basals is noticeably lower in the fourth year.

For the activity/task scales, about 25% of the planned instructional activities consisted of independent work during the first two years, dropping to about 15% in the remaining two years. Conversely, group work represented about 75% of th planned instruction during the first two years, increasing to about 85% in instructional years three and four. This is in contrast to the English data where a split of 30% versus 70% was found over all years. The relative level of formal language demand associated with this work mirrored that found in the English data: independent work about mid-level and stable across years, with group work also stable, but higher in formal demands.

For the instructor scales, as in the English data, a teacher was primarily responsible for the delivery of the planned instruction in each of the four instructional years. The planned role of the instructor tended to be mid-way between facilitation and direct instruction during the first two years, as in the English data. However, unlike the English data which showed no change in the planned role over the four instructional years, here, the role was substantially higher (reflecting almost exclusive direct instruction) for the latter two years.

Concerning the number of students expected in each group, the trend is toward larger groups (as was found in the English data) with an average of about nine students in the first years, and about 13 students in later years.

For the time measures, the average amount of planned daily reading instruction is about 30 minutes in each instructional year except the second, where it is almost 40 minutes per day (but note the large variability in that year). Thus, the amount of planned reading time for students in Spanish reading programs is about half that planned for students in English reading programs.

For the reading group rankings, the internal rank is about average across the four years, as was found in the English data. However, the external ranking is well below average in each year, and is lower here than in the English data.

Finally, the number of Checklists averaged within a semester shows the same trend found in the English data and expected from the change in data collection schedule: on average, three interviews were conducted with each teacher in each year.

The Spanish quality and quantity indices for each individual site are presented in Appendix I, again, consisting of 10 tables, two for each of the five sites with the first of a pair displaying the quality indices, and the second, the quantity indices. As with the English data, the interested reader is left to explore these differences.

Factor Analyses

In order to further reduce these instructional data, the 14 instructional quality indices and the seven associated quantity indices described above, were subjected to factor analysis techniques. The analyses were conducted over the entire set of individual summaries obtained from the bilingual sample (254 students), disregarding student, grade, and date of interview, but treating English and Spanish separately (1393 and 550 individual student summaries, respectively, obtained over the five years of data collection).

As in the RAMOS data, an examination of the descriptive statistics and correlations based on the individual protocol summary measures argued for the deletion of some of the variables from the factor analyses. Of the set of 21 indices, only 14 were entered in the English analysis, and only 10 in the Spanish analysis. The variables deleted and the rationale for their deletion are given below.

For the English analysis, a number of variables were deleted due to negligible variance. Each variable represented a quantity index (i.e., a percentage of time), with all cases having values of 100%: MAT1PR, CLSSPR, ROLEPR, IFTTPR, and ATTTPR. Unlike the RAMOS data where only one variable was deleted due low variability, such cases here are expected given the structure of the interview. A number of quantity indices were also deleted due to their strong correlations with other variables entered in the factor analysis: (1) IFWDPR, given its high correlation with IFWDRP, (2) ATLRPR, highly correlated with ATINPR, (3) IFT2PR, related to both IFWDPR and IFSTPR, (4) IFLTRP, correlated with IFLTPR, and (5) ATLRRP, related to both ATLRPR and ATINPR. Finally, a number of variables were deleted due to a relatively small number of cases for which a value was available (any value which was based on less than 80% of the total number of cases). These included (1) IFWDMN (742 cases or 53% of the sample with values), reflecting here, as in the RAMOS deletion, the little attention given to instruction in word meanings, and (2) RNKIMN (654 or 47%), the small sample size here indicating that many reading groups within a classroom were all reading in the same parts of the basal series, and therefore, no internal rankings could be made, though external rankings were possible.

For the Spanish analysis, the same variables deleted for negligible variance in the English data were deleted here (namely, MAT1PR, CLSSPR, ROLEPR, IFTTPR, and ATTTPR), as were those deleted due to their strong correlations with other variables entered in the analysis (namely, IFWDPR, ATLRPR, IFT2PR, IFLTRP, and ATLRRP). Finally, more variables were deleted in the Spanish data than in the English data due to relatively small samples, and these included the following: IFWDMN (222 or 40% of the sample), IFSTMN (307 or 56%), ATINMN (329 or 60%), RNKIMN (217 or 39%), RNKEMN (392 or 71%), and IFWDRP (406 or 74%).

English Factor Analysis

In this section, the results of the factor analysis conducted on the English individual student summaries are presented, followed by descriptive data on the subsequent computation of factor scores for the individual student yearly summaries.

Derivation of Factors

In the analysis, six factors had eigenvalues above 1, but only the initial five factors were employed in subsequent analyses. The first factor accounted for 17.5% of the variance, with the next successive factors accounting for 15.5%, 11.3%, 9.5%, and 7.7%, respectively, for a total of 61.5% explained variance (chance expectation is 36%). Of the 14 variables entered in the analysis, each loaded on at least one factor except for ATLRMN, and of these all loaded on a single factor with the exceptions of MAT1MN and IFSTPR, which loaded on two factors. Table 29 gives the factor loadings for the variables entered in the English analysis. As in the RAMOS factor tables, the rows define the entered variables, and the columns are defined by the factors, using mnemonic names which best reflect the instructional dimensions identified in the analysis. Also, as in the earlier

Table 29

Checklist - English:
Varimax Rotated Factor Matrix
based on 1393 Observations from the Bilingual Sample

		Factor				
		ADC	QFL	STW	PMT	GRV
	IFLTMN	0.023	0.639	0.064	0.285	0.007
	IFSTMN	0.194	0.026	-0.049	-0.793	-0.065
	MAT1MN	0.172	0.524	0.077	0.596	-0.129
	NBRDMN	0.029	0.196	-0.213	0.551	-0.307
	ATINMN	0.140	0.658	-0.254	0.151	-0.153
	ATLRMN	0.409	0.217	0.067	0.148	0.214
Checklist	CLSSMN	-0.155	0.657	0.165	-0.227	0.083
Variable	ROLEMN	-0.081	-0.039	-0.869	0.040	0.048
	NGTDMN	0.236	0.138	0.048	-0.055	0.681
	RNKEMN	0.598	-0.229	0.213	0.164	0.161
	IFLTPR	-0.813	-0.005	0.069	0.282	0.024
	IFSTPR	0.755	0.143	0.085	-0.195	-0.536
	ATINPR	0.070	0.009	0.853	0.010	-0.044
	IFWDRP	-0.132	-0.182	-0.161	-0.120	0.748

discussed analyses, only variables with loadings greater than .45 were used in the subsequent computation of factor scores, and thus, only these variables are discussed below.

The first factor is altogether quantitative, with a strong positive weight for the percentage of time devoted to instruction in the meaning of sentences and texts (IFSTPR), and a strong negative weight for the percentage of time devoted to decoding instruction (IFLTPR). The only other variable with a significant contribution is the external ranking of the reading group based on the students' position in the basal reader (RNKEMN), showing a positive loading. The factor resembles one of the RAMOS factors reflecting the amount of decoding. Note, however, that the loading is exactly opposite the RAMOS factor, and thus, as laid out here, this factor represents the amount of comprehension, or ACM. The positive loading of the rankin variable indicates a trend for groups above grade level to spend larger proportions of time on "reading for comprehension" while groups below grade level spend larger proportions of time on decoding.

The second factor consists entirely of positively loading quality variables: the quality of decoding instruction (IFLTMN), the quality of independent work (ATINMN), and the quality of the educational background of the individual providing instructional guidance (CLSSMN). Of lesser importance is the positive loading for primary materials (MAT1MN). Paralleling one of the RAMOS factors, this factor is best described as an index of the quality of formal language, or QFL.

The third factor is straightforward, showing a strong positive weight for the percentage of time devoted to independent work (ATINPR) and a large negative weight for the role played by any associated instructor (ROLEMN). This combination of variables indicates that the instructional role is diminished qualitatively as more reliance is given to independent work. Accordingly, this factor is labelled seatwork, or STW.

The fourth factor is complex. It shows a large positive weight for the quality of instruction devoted to sentence/text meaning (IFSTMN), but negative weights of lesser magnitude for both the quality of the primary materials (MAT1MN) and the number of basals employed (NBRDMN). This pattern suggests that greater amounts of non-explicit instruction in sentence/text meanings are found as student's are exposed to more textual material. Based on the two references to materials, this factor is labelled primary materials, or PMT, but the strong weight for IFSTMN indicates that this is an oversimplification.

The final factor derived shows a strong positive weight for the relative percentage of instruction devoted to word meaning (IFWDRP), coupled with a negative weight (of a lesser magnitude) for the percent of time devoted to sentence/text meaning (IFSTPR). The factor also shows a large positive weight for the number of students contained in the instructional group (NSTDMN). As such, the factor is best described as an index of group vocabulary instruction, or GRV, a separable component not observed in operation in the RAMOS data.

654

In summary, the five factors identified in the analysis are (1) the amount of comprehension instruction, (2) the quality of formal language, an index of the formal language demands made upon the students, (3) the amount of seatwork, (4) the quality of primary materials, although this factor is more complex, and (5) the amount of group vocabulary instruction.

Descriptive Statistics on Factor Scores

On the basis of this analysis, factor scores were created by first transforming the yearly averaged values (those final summary indices described earlier which were the averaged Fall and Spring average scale values) to z-scores, and then weighting the relevant component z-scores for each factor (those showing loadings greater than .45) by the appropriate loading values.

In Tables 30 and 31, the descriptive statistics for the computed English factor scores for the bilingual sample are presented for each site under each instructional year (Table 30 containing instructional year 0 and Table 31 displaying instructional years 1 through 4).

Note that the procedure for deriving factor scores involves standardization of the averaged summary indices first, then weighting and summing of factor component values. Thus, the resulting factor scores are not expected to conform to a standard distribution with a mean of 0 and a standard deviation of 1. As seen in Table 31, the mean of each of the (overall) factor scores at each instructional year is close to 0 (ranging from -.004 to 0.039 over the four instructional years). The standard deviations, however, show a much larger range (from 0.280 to 0.810). Although some specific site differences in factor score averages are apparent, they will not be discussed here.

In Tables 32 through 36, the correlations for the bilingual sample between the computed English factor scores and each yearly averaged summary z-score (regardless of whether or not the variable was included in the factor analysis) are presented, for instructional years 0 through 4, respectively. These correlations are discussed below, again, treating only those for instructional years 1 through 4.

For the first factor, the amount of comprehension (ACM), two variables which were not included in the analysis show positive relations with the factor in each of the four instructional years. First, the internal ranking variable, RNKIMN, shows a small relationship (about .4) which further supports the interpretation based on the loading of the external ranking variable, RNKEMN. Second, the large positive correlations for the IFT2PR variable (not included in the analysis) which represents the percentage of time devoted to instruction in meaning (word, sentence, and text combined), are expected given the high correlations with the loading variables IFLTPR and IFSTPR. Small positive relations are apparent at each instructional year (about .4) for the quality of instruction devoted to sentence/text meaning (IFSTMN), a variable included in the factor analysis. This suggests that for planned instruction, as the quantity

Table 30

Checklist - English:
Descriptive Statistics on Factur Scores Overall and by Site
for Instructional Year 0 for the Bilingual Sample

Factor	Statistic	Overall	Site 0	Site 1	Site 2	Site 3	Site 5
ACM	M	0.010	0.240	-0.630	-0.520	0.196	-0.171
	S	0.649	0.229	0.000	0.395	0.689	0.695
	N	163	32	3	12	60	56
DFL	M	0.004	0.076	0.330	0.495	-0.119	-0.027
	S	0.337	0.487	0.000	0.248	0.229	0.231
	N	163	32	3	12	60	56
STM	M	0.000	-0.606	1.110	-0.047	0.480	-0.218
	S	0.731	0.541	0.000	0.397	0.641	0.594
	N	163	32	3	12	60	56
PMT	M	-0.011	0.179	0.140	0.680	-0.222	-0.050
	S	0.473	0.382	0.000	0.398	0.418	0.424
	N	163	32	3	12	60	56
GRV	M	0.000	-0.007	-0.240	-0.235	-0.004	0.073
	S	0.487	0.246	0.000	0.284	0.523	0.578
	N	163	32	3	12	60	56

656

Table 31

Checklist - English: Descriptive Statistics on Factor Scores Overall and by Site for Instructional Years 1-4 for the Bilingual Sample

INSTRUCTIONAL YEAR 1

Factor	Statistic	Overall	Site 0	Site 1	Site 2	Site 3	Site 5
ACM	M	0.039	-0.076	-0.411	1.068	-0.199	-0.083
	S	0.733	0.328	0.245	1.186	0.454	0.398
	N	203	53	9	31	70	40
QFL	M	-0.002	-0.016	-0.121	0.011	-0.074	0.159
	S	0.324	0.337	0.210	[illegible]	0.309	0.364
	N	203	53	9	31	70	40
STW	M	0.000	-0.609	0.041	0.995	-0.158	0.307
	S	0.808	0.521	0.486	0.746	0.562	0.739
	N	203	53	9	31	70	40
PMT	M	-0.010	0.314	0.388	-0.347	-0.248	0.143
	S	0.531	0.627	0.165	0.626	0.290	0.292
	N	203	53	9	31	70	40
GRV	M	0.005	0.198	-0.004	-0.564	0.112	0.004
	S	0.447	0.476	0.058	0.285	0.375	0.305
	N	203	53	9	31	70	40

INSTRUCTIONAL YEAR 2

Factor	Statistic	Overall	Site 0	Site 1	Site 2	Site 3	Site 5
ACM	M	0.000	0.093	-0.393	-0.005	-0.049	0.077
	S	0.579	0.569	0.561	0.931	0.451	0.390
	N	228	54	13	35	75	51
QFL	M	0.002	-0.118	-0.125	-0.066	0.026	0.172
	S	0.340	0.434	0.490	0.285	0.293	0.177
	N	228	54	13	35	75	51
STW	M	0.000	-0.593	-0.811	0.724	0.046	0.269
	S	0.810	0.648	0.575	0.727	0.638	0.703
	N	228	54	13	35	75	51
PMT	M	-0.004	0.379	[illegible]	-0.137	-0.179	-0.092
	S	0.432	0.568	0.260	0.144	0.298	0.342
	N	228	54	13	35	75	51
GRV	M	[illegible]	-0.179	0.136	-0.068	0.020	0.160
	S	[illegible]	0.332	0.396	0.513	0.403	0.430
	N	228	54	13	35	75	51

INSTRUCTIONAL YEAR 3

Factor	Statistic	Overall	Site 0	Site 1	Site 2
ACM	M	0.000	0.246	-0.561	-0.146
	S	0.580	0.609	0.218	0.399
	N	97	51	14	32
QFL	M	0.000	-0.097	0.063	0.179
	S	0.294	0.262	0.212	0.322
	N	97	51	14	32
STW	M	0.000	0.348	0.266	-0.670
	S	0.806	0.622	0.520	0.759
	N	97	51	14	32
PMT	M	-0.001	0.030	0.183	-0.133
	S	0.322	0.368	0.263	0.195
	N	97	51	14	32
GRV	M	0.002	-0.277	0.428	0.262
	S	0.482	0.373	0.489	0.338
	N	97	51	14	32

INSTRUCTIONAL YEAR 4

Factor	Statistic	Overall	Site 0	Site 1	Site 2
ACM	M	0.000	0.152	-0.472	-0.057
	S	0.570	[illegible]	0.605	0.290
	N	60	[illegible]	9	17
QFL	M	0.000	0.009	-0.045	0.004
	S	0.280	0.343	0.206	0.152
	N	60	34	9	17
STW	M	0.000	-0.185	0.173	0.280
	S	0.779	0.673	0.656	0.958
	N	60	34	9	17
PMT	M	0.000	0.088	-0.281	-0.028
	S	0.381	0.409	0.396	0.226
	N	60	34	9	17
GRV	M	0.000	-0.182	0.052	0.340
	S	0.444	[illegible]	0.404	0.269
	N	60	34	9	17

657

Table 32

Checklist - English:
Correlations between Scale Z-Scores and Factor Scores
for Instructional Year 0 for the Bilingual Sample

	Entered in Analysis	Factor Scores (N=163) ACM	QFL	STW	PMT	GRV
IFLTMN	yes	0.075	0.566	-0.053	0.190	-0.077
IFWDMN	no	0.040	0.305	0.399	0.196	0.074
IFSTMN	yes	0.062	-0.246	-0.324	-0.831	0.232
MAT1MN	yes	0.111	0.588	-0.039	0.783	-0.270
NBRDMN	yes	0.161	0.323	-0.347	0.578	-0.224
ATINMN	yes	0.109	0.656	-0.109	0.447	-0.239
ATLRMN	yes	-0.065	-0.034	0.248	-0.116	0.318
CLSSMN	yes	-0.161	0.372	0.003	-0.207	0.316
ROLEMN	yes	0.021	0.052	-0.851	0.018	-0.088
NSTDMN	yes	-0.266	-0.111	0.446	-0.396	0.669
TIMEMN	no	0.339	-0.285	0.169	-0.057	-0.002
RNKIMN	no	-0.102	-0.091	0.295	0.264	0.310
RNKEMN	yes	0.621	-0.115	0.574	0.013	-0.361
NCISMN	no	0.017	0.329	-0.232	0.376	-0.179
IFLTPR	yes	-0.854	-0.102	0.238	-0.030	0.213
IFWDPR	no	0.302	0.073	-0.198	-0.045	0.399
IFSTPR	yes	0.892	0.063	-0.114	0.097	-0.778
ATINPR	yes	-0.239	-0.138	0.845	-0.005	0.219
ATLRPR	no	0.239	0.138	-0.845	0.005	-0.219
IFT2PR	no	0.854	0.102	-0.237	0.031	-0.213
IFWDRP	yes	-0.267	-0.059	-0.141	-0.283	0.808

658

Table 38

Checklist - English:
Correlations between Scale Z-Scores and Factor Scores
for Instructional Year 1 for the Bilingual Sample

	Entered in Analysis	Factor Scores (N=203)				
		ACM	QFL	STW	PNT	GRV
IFLTMN	yes	0.251	0.667	0.010	-0.074	0.086
IFWDMN	no	-0.261	-0.224	-0.097	-0.412	0.122
IFSTMN	yes	0.421	-0.209	0.260	-0.796	-0.179
MAT1MN	yes	-0.251	0.380	-0.164	0.848	-0.164
NBRDMN	yes	-0.314	0.164	-0.406	0.816	0.094
ATINMN	yes	0.070	0.504	0.049	-0.084	0.051
ATLRMN	yes	0.312	0.043	0.223	-0.138	-0.230
CLSGMN	yes	0.134	0.532	0.158	-0.001	0.000
ROLEMN	yes	-0.484	-0.050	-0.938	0.234	0.503
NSTBMN	yes	-0.163	0.037	-0.148	-0.009	0.634
TIMEMN	no	-0.292	-0.049	-0.270	-0.018	0.272
RNKIMN	no	0.305	0.097	0.038	0.219	-0.207
RNKEMN	yes	0.676	0.294	0.285	0.247	-0.277
NCISMN	no	0.207	-0.020	0.122	0.106	-0.160
IFLTPR	yes	-0.925	-0.012	-0.361	0.479	0.460
IFWDPR	no	-0.030	0.101	-0.393	-0.336	0.510
IFSTPR	yes	0.955	-0.031	0.540	-0.340	-0.692
ATINPR	yes	0.473	-0.046	0.936	-0.326	-0.413
ATLRPR	no	-0.473	0.047	-0.935	0.326	0.412
IFT2PR	no	0.925	0.012	0.361	-0.479	-0.460
IFWDRP	yes	-0.718	-0.026	-0.374	-0.266	0.724

Table 34

Checklist - English:
Correlations between Scale Z-Scores and Factor Scores
for Instructional Year 2 for the Bilingual Sample

	Entered in Analysis	Factor Scores (N=228) ACM	QFL	STW	PMT	SRV
IFLTMN	yes	0.087	0.566	0.034	-0.077	0.146
IFWDMN	no	-0.051	-0.190	-0.171	0.005	-0.270
IFSTMN	yes	0.261	0.067	0.387	-0.812	0.000
MAT1MN	yes	0.140	0.554	-0.157	0.525	-0.003
NBRDMN	yes	-0.103	-0.229	-0.209	0.647	-0.158
ATINMN	yes	0.295	0.524	0.034	-0.157	-0.004
ATLRMN	yes	0.380	0.191	-0.173	-0.010	-0.193
CLSSMN	yes	-0.010	0.552	0.081	0.056	0.039
ROLEMN	yes	-0.119	-0.060	-0.941	0.351	-0.128
NSTDMN	yes	-0.028	0.144	-0.099	-0.030	0.470
TIMEMN	no	0.138	0.294	0.410	-0.367	0.151
RNKIMN	no	0.314	0.080	0.139	-0.183	0.014
RNKEMN	yes	0.603	0.221	0.128	-0.062	-0.165
NCISMN	no	0.047	0.040	-0.320	0.589	-0.127
IFLTPR	yes	-0.886	-0.193	-0.166	0.180	0.304
IFWDPR	no	0.216	0.113	0.140	-0.270	0.462
IFSTPR	yes	0.866	0.134	0.079	0.011	-0.743
ATINPR	yes	0.176	-0. 47	0.939	-0.350	-0.096
ATLRPR	no	-0.176	0.047	-0.939	0.350	0.096
IFT2PR	no	0.885	0.193	0.166	-0.180	-0.304
IFWDRP	yes	-0.214	0.118	0.189	-0.173	0.747

Table [illegible]5

Checklist - English:
Correlations between Scale Z-Scores and Factor Scores
for Instructional Year 3 for the Bilingual Sample

	Entered in Analysis	Factor Scores (N=97)				
		ACM	QFL	STW	PMT	GRV
IFLTMN	yes	-0.037	0.266	-0.115	-0.090	-0.057
IFWDMN	no	0.062	0.193	-0.013	0.500	-0.061
IFSTMN	yes	0.413	0.202	-0.099	-0.548	-0.074
MAT1MN	yes	0.101	0.451	0.015	0.516	0.368
NBRGMN	yes	0.439	0.108	0.359	[illegible].408	-0.269
ATINMN	yes	-0.224	0.746	-0.639	-0.162	0.360
ATLRMN	yes	0.129	0.281	-0.223	-0.177	0.070
CLSSMN	yes	-0.031	0.440	-0.076	0.172	0.164
ROLEMN	yes	-0.385	0.443	-0.937	-0.239	0.419
NSTDMN	yes	-0.008	0.363	-0.387	0.014	0.612
TIMEMN	no	0.079	0.192	-0.078	0.196	0.354
RNKIMN	no	[illegible].346	0.080	0.175	0.099	-0.049
RNKEMN	yes	0.556	0.051	0.125	0.006	-0.010
NC[illegible]SMN	no	0.057	0.244	0.049	0.294	-0.078
IFLTPR	yes	-0.905	0.099	-0.417	-0.108	0.428
IFWDPR	no	0.061	0.114	0.099	0.249	0.503
IFSTPR	yes	0.889	-0.173	0.366	-0.043	-0.751
ATINPR	yes	0.357	-0.371	0.935	0.336	-0.438
ATLRPR	no	-0.357	0.371	-0.934	-0.336	0.438
IFT2PR	no	0.905	-0.099	0.417	0.108	-0.428
IFWDRP	yes	-0.381	0.382	-0.264	0.241	0.842

Table 36

Checklist - English:
Correlations between Scale Z-Scores and Factor Scores
for Instructional Year 4 for the Bilingual Sample

	Entered in Analysis	Factor Scores (N=60)				
		ACM	QFL	STW	PMT	GRW
IFLTMN	yes	-0.458	0.477	-0.111	-0.006	0.070
IFWDMN	no	0.262	-0.079	-0.277	0.091	-0.650
IFSTMN	yes	0.544	0.326	0.172	-0.804	-0.047
MATIMN	yes	-0.175	0.165	-0.008	0.642	-0.237
NBRDMN	yes	0.296	-0.087	-0.130	0.221	-0.151
ATINMN	yes	0.342	0.409	-0.381	-0.319	0.078
ATLRMN	yes	0.231	0.340	-0.397	-0.137	-0.212
CLSSMN	yes	0.257	0.701	0.146	-0.499	0.108
ROLEMN	yes	-0.029	0.100	-0.908	0.344	0.010
MSTDMN	yes	0.375	0.148	-0.198	-0.412	0.442
TIMEMN	no	0.174	0.172	-0.044	-0.360	-0.180
RNKIMN	no	0.644	0.378	0.400	-0.378	0.041
RNKEMN	yes	0.597	0.136	0.280	-0.427	0.101
NCISMN	no	0.138	-0.145	-0.091	0.118	-0.070
IFLTPR	yes	-0.890	0.055	0.021	0.259	0.141
IFWDPR	no	-0.070	-0.027	-0.028	-0.139	0.809
IFSTPR	yes	0.833	-0.027	0.002	-0.121	-0.740
ATINPR	yes	0.132	-0.272	0.904	0.009	-0.326
ATLRPR	no	-0.132	0.272	-0.904	-0.009	0.327
IFT2PR	no	0.890	-0.055	-0.021	-0.259	-0.141
IFWDRP	yes	-0.382	-0.056	-0.146	0.020	0.850

662

of this instruction in meaning increases, there is a trend for the quality to likewise increase.

For the second factor, the quality of formal language (QFL), one noticeable perturbation in the correlations is found. The variable reflecting the quality of the primary materials (MAT1MN), which was the lowest value variable to load on this factor, shows relatively lower correlations with the factor in both the first and (more particularly) the last year.

Considering the correlation pattern for the third factor, seatwork (STW), the only systematic pattern of note is that found for ATLRPR, not included in the factor analysis, which represents the percentage of time spent in group work. Given this variable's strong correlation with the percent of time in independent work (ATINPR), which loaded on this factor, the high negative correlations are to be expected.

For the fourth factor, the quality of primary materials (PMT), the only notable pattern is that for the contribution made by the number of basal texts referenced (NBRDMN), which steadily decreases with increases in instructional years (from .82 to .22).

For the final factor, group vocabulary instruction (GRV), the percentage of time devoted to instruction in word meaning (IFWDPR), which was not included in the analysis, shows the expected positive relationship with the factor, given its positive correlation with the loading variable IFWDRP.

Spanish Factor Analysis

In this section, the results of the factor analysis conducted on the Spanish individual student summaries are presented, followed by descriptive data on the subsequent computation of factor scores.

Derivation of Factors

In the analysis, six factors had eigenvalues above 1, but only the initial five factors were employed in subsequent analyses. The first factor accounted for 19.7% of the variance, with the next successive factors accounting for 15.9%, 14.4%, 11.8%, and 10.0%, respectively, for a total of 71.8% explained variance (chance expectation is 50%). Of the 10 variables entered in the analysis, each loaded on at least one factor, all loading on a single factor with the exception of NBRDMN, which loaded on two factors. Table 37 gives the factor loadings for the variables entered in the Spanish analysis. As in the English table, the rows define the entered variables, and the columns are defined by the factors, using mnemonic names which best reflect the instructional dimensions identified in the analysis. Again, only variables with loadings greater than .45 were used in the subsequent computation of factor scores, and thus, only these variables are discussed below.

Table 37

Checklist - Spanish:
Varimax Rotated Factor Matrix
based on 550 Observations from the Bilingual Sample

Checklist Variable	Factor ADC	STW	FMT	DTC	NST
IFLTMN	0.263	0.122	-0.080	-0.708	0.177
MAT1MN	0.022	-0.146	0.735	0.223	0.317
NBRDMN	0.490	0.021	0.636	-0.078	-0.103
ATLRMN	-0.286	0.326	0.586	-0.123	-0.196
CLSSMN	0.108	0.148	-0.032	0.788	0.161
ROLEMN	-0.082	-0.820	0.168	-0.164	0.046
NSTCMN	-0.051	-0.012	0.030	-0.007	0.927
IFLTPR	0.861	0.107	0.161	0.102	0.133
IFSTPR	-0.776	-0.029	0.172	0.254	0.212
ATINPR	0.055	0.835	0.207	-0.111	0.029

664

The first factor is similar to a factor found in the English data; it reflects the amount of decoding instruction with a strong positive weight for IFLTPR and a large negative weight for IFSTPR. A final weakly loading variable is the positive one for the number of basal texts used (NBRDMN). Though the factor parallels the English ACM factor, the weighting is in the opposite direction, and thus is labelled amount of decoding, or ADC.

The second factor is quite similar to the third English Checklist factor reflecting the amount of seatwork. As found in those data, there is a positive weighting for the proportion of time devoted to independent work (ATINPR), coupled with a strong negative weight for the role played by any associated instructor (ROLEMN). As before, this factor is labelled seatwork, or STW.

The third factor appears to be group instruction centered around the use of the basal. It shows a large positive weight for the quality of the primary materials (MAT1MN), and smaller positive weights for the number of basal texts employed (NBRDMN) and the quality of formal language associated with group work. Based on the two weights for materials, the factor is labelled primary materials, or PMT, although this is oversimplified.

The fourth factor appears to be contrastive with the first factor, the amount of decoding instruction. This factor shows a large negative weight for the quality of instruction devoted to decoding (IFLTMN) and a large positive weight for the classification of any associated instructor (CLSSMN). This pattern suggests that the planned decoding instruction supervised by aides is more explicit than that to be supervised by teachers -- aides are employed to teach letter-sound correspondences, while teachers provide less explicit practice (to those who perhaps have already acquired the rudiments of decoding). Accordingly, this factor is labelled decoding teacher classification, or DTC.

The fifth factor is straightforward, showing a single loading for the number of students in the instructional group. As such, this factor is labelled number of students, or NST.

In summary, the five factors identified in the analysis are (1) the amount of decoding instruction, (2) the amount of seatwork, (3) the quality of primary materials, although this factor is more complex, (4) the decoding teacher's classification, which is associated with the explicitness of the decoding instruction planned, and (5) the number of students in the instructional group.

Descriptive Statistics on Factor Scores

On the basis of this analysis, factor scores were created following the same procedure used in the computation of factor scores for the English summaries: first transforming the yearly averaged values to z-scores, and then weighting the relevant component z-scores for each factor (those showing loadings greater than .45) by the appropriate loading values.

665

In Tables 38 and 39, the descriptive statistics for the computed Spanish factor scores for the bilingual sample are presented for each site under each instructional year (Table 38 containing instructional year 0 and Table 39 displaying instructional years 1 through 4).

Again, the resulting factor scores are not expected to conform to a standard distribution. As seen in Table 39, the mean of each of the (overall) factor scores at each instructional year is close to 0 (ranging from -0.002 to 0.007). As in the English data, the standard deviations show a much larger range (from 0.248 to 1.001). Although some specific site differences in factor score averages are apparent, they will not be discussed here.

In Tables 40 through 44, the correlations for the bilingual sample between the computed Spanish factor scores and each yearly averaged summary z-score are presented for instructional years 0 through 4, respectively. Before discussing these, note that as in the corresponding Spanish RAMOS tables, the coefficients for instructional year 4 (Table 44) are uniformly high. These generally reflect the reduced sample size (particularly, the reduced number of classrooms offering Spanish reading), and thus, interpretations of the planned instruction within this year must be treated cautiously.

For the first factor, the amount of decoding instruction (ADC), the lowest loading variable was the number of basal texts referenced (NBRDMN), and the correlations of this variable with the factor score show similarly low positive coefficients (about .35). Interestingly, both of the rank variables, RNKIMN and RNKEMN, show generally negative correlations with the factor score over the instructional years, and these indicate that increased percentages of time are planned in decoding instruction for the lower reading groups. Finally, the consistently high negative correlations of IFT2PR (not included in the analysis) with the factor over the instructional years are expected given the high correlations with IFLTPR and IFSTPR, both loading on this factor.

For the second factor, seatwork (STW), the quality of instruction in word meaning (IFWDMN) which was not included in the analysis shows consistently negative correlations (about -.45) with the factor. This suggests the trend for high quality instruction in word meaning to be seldom planned as seatwork. Given the compilmentary relationship between group work and independent work, the factor's high negative correlations with ATLRMN across the instructional years are expected.

The correlation pattern for the third factor, the quality of primary materials (PMT), shows the expected pattern given the relative strength of the loadings: higher correlations for the two highest loading variables (MATIMN and NBRDMN) and lower correlations for the lowest loading variable (ATLRMN).

The fourth factor, decoding teacher classification (DTC), similarly shows the pattern expected given the factor loadings, with no other correlations (from the non-loading varibles) above .4 in the initial instructional years.

Table 38

Checklist - Spanish:
Descriptive Statistics on Factor Scores Overall and by Site
for Instructional Year 0 for the Bilingual Sample

Factor	Statistic	Overall	Site 0	Site 1	Site 2	Site 3	Site 5
ADC	M	0.000	0.286	-0.040	0.666	-0.840	-0.037
	S	0.596	0.250	0.000	0.111	0.470	0.556
	N	93	26	4	9	14	40
STW	M	0.000	-0.052	1.230	0.046	0.037	-0.111
	S	0.564	0.534	0.000	0.121	0.476	0.559
	N	93	26	4	9	14	40
PMT	M	0.000	0.146	-0.290	-0.186	-0.311	0.085
	S	0.390	0.253	0.000	0.519	0.260	0.408
	N	93	26	4	9	14	40
OTC	M	0.004	-0.819	0.040	0.672	0.465	0.223
	S	0.653	0.590	0.000	0.279	0.343	0.216
	N	93	26	4	9	14	40
NST	M	0.000	0.173	1.630	-0.520	-0.018	-0.150
	S	1.000	0.761	0.000	0.000	1.293	1.026
	N	93	26	4	9	14	40

667

Table 39

Checklist - Spanish:
Descriptive Statistics on Factor Scores Overall and by Site
for Instructional Years 1-4 for the Bilingual Sample

INSTRUCTIONAL YEAR 1

Factor	Statistic	Overall	Site 0	Site 1	Site 2	Site 3	Site 5
ADC	M	0.000	-0.041	-0.566	0.503	-0.606	0.046
	S	0.578	0.475	0.380	0.158	0.373	0.609
	N	75	17	8	15	6	29
STW	M	0.000	-0.714	-0.127	0.043	-0.365	0.506
	S	0.683	0.424	0.235	0.169	0.725	0.631
	N	75	17	8	15	6	29
PMT	M	0.000	0.070	-0.468	0.343	-0.848	0.087
	S	0.422	0.254	0.573	0.167	0.139	0.207
	N	75	17	8	15	6	29
DTC	M	0.000	-0.285	0.195	0.358	0.118	-0.096
	S	0.602	0.923	0.275	0.195	0.532	0.489
	N	75	17	8	15	6	29
NST	M	0.000	-0.090	0.632	0.875	-0.630	-0.441
	S	0.999	0.517	0.686	1.678	0.200	0.377
	N	75	17	8	15	6	29

INSTRUCTIONAL YEAR 2

Factor	Statistic	Overall	Site 0	Site 1	Site 2	Site 3	Site 5
ADC	M	0.000	0.347	-0.160	-0.354	-0.012	-0.285
	S	0.520	0.358	0.361	0.341	0.804	0.478
	N	90	38	8	10	5	29
STW	M	0.000	-0.320	-0.585	0.452	0.712	0.303
	S	0.784	0.860	0.649	0.239	1.188	0.392
	N	90	38	8	10	5	29
PMT	M	0.000	0.107	-0.213	-0.365	-0.030	0.050
	S	0.419	0.448	0.721	0.349	0.261	0.188
	N	90	38	8	10	5	29
DTC	M	0.000	-0.210	0.226	0.590	-0.492	0.095
	S	0.565	0.632	0.348	0.000	0.754	0.354
	N	90	38	8	10	5	29
NST	M	0.000	-0.217	-0.026	-0.186	[illegible]	0.432
	S	1.000	0.579	0.461	0.285	[illegible]	1.334
	N	90	38	8	10	5	29

INSTRUCTIONAL YEAR 3

Factor	Statistic	Overall	Site 0	Site 1	Site 2
ADC	M	0.000	0.075	0.036	-0.274
	S	0.532	0.586	0.268	0.244
	N	48	35	3	10
STW	M	0.000	0.069	1.076	-0.566
	S	0.804	0.800	0.529	0.366
	N	48	35	3	10
PMT	M	0.001	-0.088	-0.234	0.390
	S	0.305	0.513	0.583	0.210
	N	48	35	3	10
DTC	M	0.000	-0.016	-0.153	0.102
	S	0.504	0.577	0.457	0.025
	N	48	35	3	10
NST	M	0.000	-0.454	-0.166	1.640
	S	1.001	0.421	1.261	0.638
	N	48	35	3	10

INSTRUCTIONAL YEAR 4

Factor	Statistic	Overall	Site 0	Site 1
ADC	M	0.000	-0.061	0.580
	S	0.542	0.534	0.000
	N	21	19	2
STW	M	0.000	0.002	-0.020
	S	0.679	0.716	0.000
	N	21	19	2
PMT	M	0.001	0.013	-0.110
	S	0.248	0.258	0.000
	N	21	19	2
DTC	M	0.007	0.024	-0.160
	S	0.379	0.396	0.000
	N	21	19	2
NST	M	-0.002	-0.207	1.950
	S	1.000	0.802	0.000
	N	21	19	2

668

Table 40

Checklist - Spanish:
Correlations between Scale Z-Scores and Factor Scores
for Instructional Year 0 for the Bilingual Sample

	Entered in Analysis	Factor Scores (N=93)				
		ADC	STW	PMT	DTC	NST
IFLTMN	yes	0.198	0.082	0.292	-0.863	0.372
IFWDMN	no	-	-	-	-	-
IFSTMN	no	-0.016	0.540	-0.374	-0.076	-0.128
MATIMN	yes	0.047	-0.148	0.730	0.131	-0.158
NBRDMN	yes	0.817	-0.336	0.546	-0.347	-0.152
ATINMN	no	0.551	-0.150	0.431	-0.567	-0.195
ATLRMN	yes	-0.184	0.116	0.486	-0.131	0.127
CLSSMN	yes	-0.159	0.206	-0.029	0.899	-0.111
ROLEMN	yes	0.375	-0.674	0.369	-0.377	0.192
NSTDMN	yes	-0.063	-0.060	-0.119	-0.245	1.000
TIMEMN	no	-0.384	0.167	-0.068	0.244	0.220
RNKIMN	no	0.553	-0.875	0.765	-0.868	-0.927
RNKEMN	no	0.245	0.109	0.276	-0.596	-0.329
NCISMN	no	0.406	0.436	0.151	-0.176	0.171
IFLTPR	yes	0.843	0.172	0.280	0.148	-0.054
IFWDPR	no	-0.453	-0.263	-0.186	-0.397	0.068
IFSTPR	yes	-0.852	0.107	-0.224	0.374	-0.011
ATINPR	yes	0.293	0.688	0.068	-0.256	0.080
ATLRPR	no	-0.293	-0.688	-0.068	0.256	-0.080
IFT2PR	no	-0.844	-0.172	-0.280	-0.147	0.053
IFWDRP	no	0.233	-0.417	-0.063	-0.544	-0.036

669

Table 41

Checklist - Spanish:
Correlations between Scale Z-Scores and Factor Scores
for Instructional Year 1 for the Bilingual Sample

	Entered in Analysis	Factor Scores (N=75) ADC	STW	PMT	DTC	NST
IFLTMN	yes	0.048	-0.081	-0.155	-0.781	0.006
IFWDMN	no	0.378	-0.415	-0.206	-0.283	0.033
IFSTMN	no	0.112	-0.004	0.070	-0.074	0.363
MATIMN	yes	0.136	-0.442	0.702	-0.017	0.073
NBRDMN	yes	0.450	0.198	0.852	-0.035	-0.039
ATINMN	no	-0.156	0.180	0.346	-0.132	-0.357
ATLRMN	yes	0.266	0.552	0.351	0.394	-0.064
CLSSMN	yes	-0.129	0.005	0.098	0.826	0.127
ROLEMN	yes	-0.163	-0.820	-0.214	0.031	0.287
NSTDMN	yes	0.067	-0.088	-0.007	0.079	1.000
TIMEMN	no	0.047	0.294	-0.088	0.052	0.726
RNKIMN	no	0.357	0.095	0.286	-0.017	-0.046
RNKEMN	no	-0.077	0.191	0.009	-0.233	0.209
NCISMN	no	0.122	-0.006	0.332	0.331	-0.057
IFLTPR	yes	0.945	0.088	0.230	-0.119	0.139
IFWDPR	no	-0.426	-0.014	-0.187	0.078	-0.233
IFSTPR	yes	-0.901	-0.100	-0.165	0.097	-0.020
ATINPR	yes	0.075	0.827	-0.047	0.115	0.137
ATLRPR	no	-0.076	-0.827	0.046	-0.115	-0.137
IFTZPR	no	-0.945	-0.088	-0.230	0.119	-0.139
IFWORP	no	-0.117	0.074	-0.266	0.088	-0.323

670

Table 42

Checklist - Spanish:
Correlations between Scale Z-Scores and Factor Scores
for Instructional Year 2 for the Bilingual Sample

	Entered in Analysis	Factor Scores (N=90) ADC	STW	PMT	DTC	NST
IFLTMN	yes	0.249	-0.192	0.202	-0.724	-0.121
IFWDMN	no	0.540	-0.498	0.014	-0.411	-0.318
IFSTMN	no	-0.023	0.345	0.178	0.182	0.260
MAT1MN	yes	-0.103	-0.101	0.731	-0.217	-0.083
NBRDMN	yes	0.245	-0.041	0.684	-0.026	0.064
ATINMN	no	0.013	-0.287	0.277	-0.098	0.051
ATLRMN	yes	0.107	0.194	0.489	0.056	-0.203
CLSSMN	yes	-0.229	0.320	0.018	0.784	0.106
ROLEMN	yes	0.030	-0.945	-0.052	-0.305	-0.260
NSTDMN	yes	-0.188	0.228	-0.111	0.150	1.000
TIMEMN	no	-0.434	0.231	-0.032	0.056	0.345
RNKIMN	no	-0.399	0.304	0.274	0.242	0.388
RNKEMN	no	-0.400	0.340	0.304	0.128	0.224
NCISMN	no	0.370	0.088	0.245	0.075	-0.146
IFLTPR	yes	0.854	0.077	-0.174	-0.208	-0.155
IFWDPR	no	-0.053	-0.378	0.229	-0.193	-0.091
IFSTPR	yes	-0.908	0.218	0.008	0.387	0.247
ATINPR	yes	-0.118	0.947	-0.031	0.344	0.173
ATLRPR	no	0.118	-0.948	0.032	-0.345	-0.173
IFT2PR	no	-0.853	-0.077	0.174	0.207	0.155
IFWDRP	no	0.321	-0.293	0.194	-0.292	-0.174

671

Table 43

Checklist - Spanish:
Correlations between Scale Z-Scores and Factor Scores
for Instructional Year 3 for the Bilingual Sample

	Entered in Analysis	Factor Scores (N=48) ADC	STW	PMT	DTC	NST
IFLTMN	yes	-0.310	-0.129	-0.501	-0.626	-0.015
IFWDMN	no	-0.210	-0.545	-0.247	-0.217	0.045
IFSTMN	no	0.063	0.172	0.339	-0.099	0.713
MATIMN	yes	0.265	-0.119	0.724	0.362	0.366
NBRDMN	yes	0.371	0.516	0.840	0.213	-0.047
ATINMN	no	0.000	-0.113	0.550	0.528	-0.256
ATLRMN	yes	0.053	0.299	0.769	0.183	-0.011
CLSSMN	yes	0.067	-0.244	-0.021	0.715	-0.001
ROLEMN	yes	-0.234	-0.971	-0.223	0.076	0.384
NSTDMN	yes	-0.223	-0.404	0.153	0.009	1.000
TIMEMN	no	-0.065	0.449	0.103	-0.206	0.036
RNKIMN	no	-0.415	-0.038	-0.519	-0.306	0.464
RNKEMN	no	-0.431	-0.124	0.494	0.267	0.269
NCISMN	no	0.245	0.074	-0.130	0.478	-0.169
IFLTPR	yes	0.792	-0.076	-0.141	0.127	0.002
IFWDPR	no	-0.042	0.566	0.447	0.125	-0.449
IFSTPR	yes	-0.945	-0.450	-0.254	-0.280	0.431
ATINPR	yes	0.416	0.972	0.308	-0.116	-0.400
ATLRPR	no	-0.415	-0.972	-0.307	0.117	0.399
IFT2PR	no	-0.792	0.077	0.142	-0.127	-0.001
IFWDRP	no	0.344	0.716	0.454	0.034	-0.621

672

Table 44

Checklist - Spanish:
Correlations between Scale Z-Scores and Factor Scores
for Instructional Year 4 for the Bilingual Sample

	Entered in Analysis	Factor Scores (N=21) ADC	STW	PMT	DTC	NST
IFLTMN	yes	0.669	-0.564	-0.538	-0.401	0.523
IFWDMN	no	-	-	-	-	-
IFSTMN	no	0.264	0.863	0.655	0.423	-0.382
MAT1MN	yes	0.905	0.481	0.785	-0.263	0.110
NBRDMN	yes	0.389	0.778	0.607	0.276	-0.267
ATINMN	no	-0.245	-0.801	-0.623	-0.074	0.457
ATLRMN	yes	-0.670	-0.409	-0.381	0.240	-0.283
CLSSMN	yes	0.037	-0.230	-0.036	0.598	0.445
ROLEMN	yes	-0.329	-0.817	-0.608	-0.340	0.276
NSTDMN	yes	0.189	-0.552	-0.344	-0.019	1.000
TIMEMN	no	0.511	0.717	0.937	0.299	-0.436
RNKIMN	no	-0.914	0.027	-0.270	0.247	0.211
RNKEMN	no	-0.103	-0.579	-0.555	-0.206	0.288
NCISMN	no	0.253	-0.419	-0.261	-0.300	0.586
IFLTPR	yes	0.797	-0.120	0.230	-0.493	0.550
IFWDPR	no	0.088	0.8[illegible]	0.674	0.255	-0.771
IFSTPR	yes	-0.958	-0.564	-0.828	0.328	0.046
ATIMPR	yes	0.410	0.825	0.743	-0.092	-0.627
ATLRPR	no	-0.411	-0.823	-0.742	0.093	0.626
IFT2PR	no	-0.798	0.119	-0.231	0.492	-0.550
IFWDRP	no	0.211	0.728	0.652	0.065	-0.767

The final factor, the number of students contained in the instructional group (NST), had a single loading variable, which shows the expected perfect correlation with the resulting factor score. Further, no other variable shows correlations above the .4 level in instructional years 1 and 2.

Correlations Between Factors

Having treated the correlations between the summary values and the individual factor scores for both the English and Spanish data, next the correlations between the factor scores themselves within each instructional year are considered. Again, the correlations for English and Spanish are discussed separately, treating English first.

English Factor Score Correlations

In this section, the correlations between the English factor scores for the bilingual sample are discussed. In Table 45, these correlation coefficients are presented for each of the five instructional years.

Considering the correlation pattern within the last four instructional years, the only systematic relationship across years is a negative one, averaging about -.5 between the amount of comprehension (ACM) and group vocabulary instruction (GRV). This relationship largely reflects the common contribution of IFSTPR, the amount of instruction devoted to sentence/text meaning, which loaded positively on the ACM factor and negatively on the GRV factor. Beyond this correlation, the most notable remaining relationships are within the first and third instructional years.

In these two instructional years, the amount of comprehension (ACM) shows a moderate positive relationship (about .45) to the amount of seatwork (STW). Further, the amount of seatwork is equally negatively related in these years to the group vocabulary factor (GRV). These relationships suggest that for students given relatively larger amounts of seatwork, there is an increase in the amount of comprehension instruction and a decrease in the amount of group vocabulary instruction.

Spanish Factor Score Correlations

The correlations between the Spanish factor scores for the bilingual sample are presented in Table 46; again, these coefficients are presented separately for each of the five instructional years, but the discussion below only treats those for the last four instructional years.

For the last four instructional years, only one relationship across the years appears to be systematic: the amount of decoding (ADC) tends to be positively related (about .4, but much less in year 2 and much more in year 4) to the primary materials factor (PMT). This most likely reflects the common positive loading of the number of

Table 45

Checklist - English:
Correlatons between Factor Scores for Each Instructional Year
for the Bilingual Sample

		Factor				
		ACM	QFL	STW	PMT	GRV
	ACM	-	0.063	-0.152	0.104	-0.578
INSTRUCTIONAL	QFL		-	-0.111	0.499	-0.102
YEAR 0	STW			-	-0.014	0.180
(N=163)	PMT				-	-0.361
	GRV					-
	ACM	-	0.025	0.511	-0.387	-0.588
INSTRUCTIONAL	QFL		-	0.002	0.298	0.020
YEAR 1	STW			-	-0.298	-0.489
(N=203)	PMT				-	-0.007
	GRV					-
	ACM	-	0.225	0.157	-0.101	-0.522
INSTRUCTIONAL	QFL		-	0.007	0.115	0.068
YEAR 2	STW			-	-0.373	0.018
(N=228)	PMT				-	-0.122
	GRV					-
	ACM	-	-0.104	0.396	0.034	-0.529
INSTRUCTIONAL	QFL		-	-0.435	0.195	0.419
YEAR 3	STW			-	0.307	-0.458
(N=97)	PMT				-	0.134
	GRV					-
	ACM	-	0.009	0.088	-0.326	-0.358
INSTRUCTIONAL	QFL		-	-0.204	-0.182	0.038
YEAR 4	STW			-	-0.186	-0.184
(N=60)	PMT				-	-0.150
	GRV					-

675

Table 46

Checklist - Spanish:
Correlatons between Factor Scores for Each Instructional Year
for the Bilingual Sample

		Factor				
		ADC	STW	PMT	DTC	NST
INSTRUCTIONAL YEAR 0 (N=93)	ADC	-	-0.056	0.382	-0.186	-0.063
	STW		-	-0.217	0.084	-0.080
	PMT			-	-0.172	-0.119
	DTC				-	-0.245
	NST					-
INSTRUCTIONAL YEAR 1 (N=75)	ADC	-	0.144	0.429	-0.112	0.067
	STW		-	0.099	0.051	-0.088
	PMT			-	0.155	-0.007
	DTC				-	0.079
	NST					-
INSTRUCTIONAL YEAR 2 (N=90)	ADC	-	-0.079	0.113	-0.315	-0.188
	STW		-	0.010	0.343	0.228
	PMT			-	-0.1[illegible]4	-0.111
	DTC				-	0.150
	NST					-
INSTRUCTIONAL YEAR 3 (N=48)	ADC	-	0.335	0.305	0.270	-0.223
	STW		-	0.274	-0.099	-0.404
	PMT			-	0.335	0.153
	DTC				-	0.009
	NST					-
INSTRUCTIONAL YEAR 4 (N=21)	ADC	-	0.451	0.702	-0.335	0.189
	STW		-	0.823	0.148	-0.552
	PMT			-	0.166	-0.344
	DTC				-	-0.019
	NST					-

676

basal readers (NBRDMN) in both factors. The only other notable trend concerns the negative relationship between the number of students factor (NST) and the amount of seatwork (STW), which would be expected given that the former tends to represent increased group work while the latter represents increased independent work.

This concludes the discussion of the Checklist data collected in English and Spanish reading classrooms. As was true in the RAMOS data, the structure of these two data sets was quite similar. In the next section, the pictures of reading instruction (in both English and Spanish) revealed by the observation and interview data will be contrasted.

RELATIONSHIPS BETWEEN RAMOS AND CHECKLIST INDICES

Although not of primary interest in this study, the relationships between the instructional indices derived from the RAMOS and Checklist data are discussed below. First, contrasts based on the summary measure descriptive data (aggregated) are discussed, followed by a presentation of the correlations between the factor scores derived from the two data sets. Within each of the presentations, the English data are given first, then the Spanish data.

Contrasts between RAMOS and Checklist Summary Measures

In this section, observed instruction, as quantified in the RAMOS summary measures, is contrasted with planned instruction, as quantified in the Checklist summary measures, both sets of measures having been individually discussed previously.

English Contrasts

In treating the English contrasts, the relevant RAMOS data come from Tables 3 and 4, the Checklist data from Tables 27 and 28. The aggregated contrasts over the final four instructional years, based on the common summary indices represented in the two instruments, can be summarized as follows:

For the number of students, the trend toward increasing group size was found under both assessments, though the Checklist data tended to underestimate what was actually observed.

Considering the instructor variables, both assessments found the teacher with primary responsibility for the delivery of instruction, though the role played tended to be overestimated in the Checklist data.

For the instructional focus variables, the same trend was found in both assessments with respect to the quality indices. However, slightly more decoding instruction was planned than observed, and, concomitantly, slightly less instruction on connected text was planned than observed.

For the activity/task scales, again, the same trend was generally found in both assessments with respect to the quality indices; however, more independent work (and thus, less group work) was observed than planned.

For the materials, the trend toward more usage of text with increasing grade levels was found in both instruments.

In general, these aggregate contrasts suggest that the instructional picture provided by the RAMOS data is not substantially different from that provided by the Checklist data. The most noticeable discrepancies concern the estimate of the instructor's role and the relative proportions of time devoted to particular instructional foci and activity/tasks; the most noticeable similarities are those related to the quality estimates of instruction (in particular, for the instructional foci and activity/tasks).

Spanish Contrasts

In treating the Spanish contrasts, the relevant RAMOS data come from Tables 5 and 6, the Checklist data from Tables 27 and 28. The aggregate contrasts over the final four instructional years, again based on the common summary indices represented in the two instruments, are summarized below following the format of the English contrasts given above:

For the number of students, the trend toward increasing group size was found under both assessments, though the Checklist data, as was found in the English data, tended to underestimate the average group size actually observed.

Considering the instructor variables, both assessments found the teacher with primary responsibility. Again, as in the English data, the role played tended to be overestimated in the Checklist data, substantially so with respect to the last two years of instruction.

For the instructional focus variables, the same trend was found in both assessments with respect to the quality indices. However, more decoding instruction was planned than observed (substantially more in the latter two years), and, concomitantly, less instruction on connected text was planned than observed.

For the activity/task scales, again, the same trend was generally found in both assessments with respect to the quality indices; however, more independent work (and thus, less group work) was observed than planned.

For the materials, the trend toward more usage of text with increasing grade levels was found in both instruments (as was found in the English data).

Thus, as in the English data, the aggregate contrasts suggest that the instructional pictures provided by the two data sets are similar, again, with greater discrepancies found in the quantity indices than in the quality indices. A better picture of the correspondence between the data sets (though not the best assessment) can be found in the pattern of factor score intra-correlations, which are discussed next.

Correlations between RAMOS and Checklist Factor Scores

The correlations between the RAMOS and Checklist factor scores by instructional year are presented in this section, treating the English based set first, then those from the Spanish data. These are presented for two reasons -- first, as a way of assessing the independence of the instructional factors derived from the two instruments given their predictive use in the analyses presented in the next volume linking reading growth and instruction, and second, as a way of assessing the degree to which instructional descriptions derived from the two instruments are similar. For the latter, correlations between the individual summary indices within each semester would provide a better assessment of this correspondence, but the relationships between relevant factors can provide some information on this issue. As an aid to the discussion, Table 47 is presented as a summary of the factors derived from these two major instruction data sets, giving the acronyms and labels for each factor.

English Correlations

Table 48 presents the RAMOS and Checklist correlations for the English data sets for each instructional year, 0 through 4. Given the emphasis on the last four instructional years, only these will be discussed here.

Concerning the independence of the indices, note that in general, the coefficients are low -- indeed, of the 140 correlations displayed in the final instructional years, only nine exceed a level of 15% shared variance, and most of these are expected (as discussed below). Thus, the factor scores provide relatively independent sources of instructional information whenever the overlap of common component variables within the two instruments is minimal.

In assessing the degree of correspondence between the two data sets using correlations between the component factor scores, two points are important to note. First, RAMOS and Checklist data for a given student do not necessarily correspond to similar points in time. Further, the assessment of this correspondence between instruments in this analysis is hampered by the averaging procedures employed within and across semesters. Thus, the magnitude of correlations expected here are difficult to gauge. Second, individual factor scores rarely contained the same component variables. The degree to which such components are shared within factors from the different instruments (and the degree of similarity in the contribution to their

Table 47

Instruction:
Factor Names and Labels

Instrument/ Language	Factor Number	Acronym	Label
RAMOS-E	1	ETT	Engaged Text Time
	2	DGI	Direct Group Instruction
	3	QFL	Quality of Formal Language
	4	ADC	Amount of Decoding
	5	PRD	Productivity
	6	SMT	Secondary Materials
	7	NST	Number of Students
RAMOS-S	1	QFL	Quality of Formal Language
	2	DGI	Direct Group Instruction
	3	ETT	Engaged Text Time
	4	NST	Number of Students
	5	ADC	Amount of Decoding
	6	SMT	Secondary Materials
	7	CNT	Control
CHECKLIST-E	1	ACM	Amount of Comprehension
	2	QFL	Quality of Formal Language
	3	STW	Seatwork
	4	PMT	Primary Materials
	5	GRV	Group Vocabulary Instruction
CHECKLIST-S	1	ADC	Amount of Decoding
	2	STW	Seatwork
	3	PMT	Primary Materials
	4	DTC	Decoding Teacher Classification
	5	NST	Number of Students

6 30

Table 48

Instruction - English:
Correlations between the Checklist
and RAMOS Factor Scores for the Bilingual Sample

Instructional Year 0
(N=134)
Checklist

		ACM	QFL	STW	PMT	GRV
	ETT	0.072	0.060	0.103	-0.045	-0.119
	DGI	-0.110	0.206	0.162	0.293	0.227
	QFL	0.033	0.100	0.194	0.226	-0.053
RAMOS	ADC	-0.018	-0.125	0.188	-0.109	-0.220
	PRD	0.136	-0.263	0.341	-0.090	0.155
	SMT	0.175	-0.001	-0.241	0.223	0.061
	NST	-0.233	-0.013	-0.305	-0.010	0.277

Instructional Year 1
(N=192)
Checklist

		ACM	QFL	STW	PMT	GRV
	ETT	-0.220	-0.199	-0.145	-0.178	0.031
	DGI	0.174	0.040	-0.319	0.198	0.089
	QFL	-0.168	0.072	-0.361	0.247	0.287
RAMOS	ADC	-0.001	0.219	-0.161	0.387	0.031
	PRD	0.080	0.034	-0.274	0.407	-0.060
	SMT	-0.065	-0.033	-0.191	0.285	0.008
	NST	0.124	0.167	0.518	-0.009	-0.287

Instructional Year 3
(N=94)
Checklist

	ACM	QFL	STW	PMT	GRV
ETT	0.107	0.080	-0.130	-0.042	-0.035
DGI	0.290	0.190	0.021	0.267	-0.001
QFL	0.359	-0.246	0.409	0.086	-0.400
ADC	-0.180	0.068	-0.060	-0.075	0.250
PRD	0.208	-0.039	0.015	-0.070	-0.151
SMT	-0.097	0.429	0.015	-0.008	0.035
NST	-0.173	0.076	-0.148	0.086	0.374

Instructional Year 2
(N=218)
Checklist

		ACM	QFL	STW	PMT	GRV
	ETT	0.013	0.195	0.243	-0.243	0.164
	DGI	-0.209	-0.309	-0.249	0.345	-0.029
	QFL	-0.072	0.003	0.029	0.057	0.025
RAMOS	ADC	0.172	0.168	-0.024	-0.018	-0.083
	PRD	0.280	-0.070	-0.320	0.442	-0.162
	SMT	-0.027	-0.026	-0.005	0.182	-0.016
	NST	-0.130	0.220	0.200	-0.364	0.342

Instructional Year 4
(N=59)
Checklist

	ACM	QFL	STW	PMT	GRV
ETT	-0.026	-0.022	-0.067	-0.042	-0.018
DGI	-0.011	0.329	-0.415	-0.180	0.178
QFL	-0.292	-0.296	-0.212	0.219	0.073
ADC	-0.200	0.217	0.026	-0.186	-0.133
PRD	0.323	0.566	-0.046	-0.553	0.026
SMT	-0.072	-0.306	0.123	0.240	-0.004
NST	-0.008	0.20[illegible]	-0.374	-0.298	0.292

681

respective factor score as determined by their relative loadings) influences the degree to which the factors will be related, but again, the magnitude of the relationships expected are difficult to anticipate. With these points in mind, the discussion below centers on the relationships between factors with common components.

First, the RAMOS factor reflecting the amount of decoding instruction (ADC) and the Checklist factor reflecting the amount of comprehension instruction (ACM) both revealed loadings for three variables, with two in common, namely IFLTPR and IFSTPR (the relative percentage of time devoted to decoding instruction and to instruction in sentence/text meaning, respectively). Given that the loadings of these variables within the two factors were opposite in sign, a negative correlation between the factors would be expected. Although the coefficients are small such negative values are found in the last two years, with a zero correlation in the first year and a small positive correlation in the second year.

Second, of the four variables loading on the quality of formal language factor (QFL) within each instrument, three were shared (IFLTMN, MAT1MN, and ATINMN). All load in the same direction, but the materials variable is the highest loading variable on the RAMOS factor and the lowest loading on the Checklist factor. Further, the highest loading variable on the Checklist factor, CLSSMN, did not load on the RAMOS factor, all leading to the expectation of a small positive correlation. From Table 48, the correlation coefficients for these factors are zero in the first two instructional years, and slightly negative in the last two instructional years (about -.27), thus going against the expected trend.

Third, the two variables loading on the seatwork factor (STW) found in the Checklist data, namely, ROLEMN and ATINPR, also loaded on the direct group instruction factor (DGI) of the RAMOS data, though four additional variables were components of this later factor. Given the opposite sign of the loadings, a negative correlation between these factors would be expected, though small due to the lesser influence of the two common variables in the RAMOS factor. As seen in Table 48, except for the zero correlation found in the third year, negative values are found in each of the remaining years (averaging about -.3).

Fourth, the group vocabulary factor (GRV) derived from the Checklist data contained three component variables, one of which was the number of students (NSTDMN). In the RAMOS data, this variable alone constituted a number of students factor (NST), and a postive relationship between these factors would be expected. From Table 48, positive relationships are found in the last three instructional years (about .3), with a negative relationship in the first year (about -.3)

Finally, the remaining two variables loading on the group vocabulary factor (GRV) derived from the Checklist data, namely, IFSTPR and IFWDRP, also loaded on the amount of decoding factor (ADC) derived from the RAMOS data. The former variable (IFSTPR) was the highest loading variable on the RAMOS factor and the lowest loading variable on the Checklist factor, while IFWDRP had the opposite pattern. Given that the loadings were of the same sign, this would lead to the expectation of a small positive relationship between the factors. The corresponding coefficients found in Table 48 are all close to zero with the exception of a positive value in the third year (.25).

In summary, although the assessment is weak (for reasons explained above), the correlation pattern between factors with shared components gives some support to a common instructional picture being provided by the two instruments.

Spanish Correlations

Table 49 presents the Spanish RAMOS and Checklist correlations for instructional years 0 through 4. Concerning the independence of the indices, note that the magnitudes of the correlations in general are higher in the last two instructional years than in the first two, and for the former, they are higher in the fourth than in the third instructional year. Recalling that the number of classrooms found to be offering Spanish reading programs was greatly reduced in these years, and that for those students in Spanish programs, the variability between individual summaries was minimal, the interpretation of these correlations is difficult. Disregarding the last two instructional years, the remaining coefficients are low -- of the 70 correlations, 8 exceed a level of 15% shared variance, again arguing that the factor scores provide relatively independent sources of instructional information whenever the overlap of common component variables within the two instruments is minimal.

In treating the correlation pattern between factors, the discussion will again center on the relationships between factors with common components. First, two variables were found to load on the amount of decoding (ADC) factor derived from the RAMOS data, namely, IFLTPR and IFSTPR. These same two variables were components of the corresponding Checklist factor along with a third weakly loading variable, NBRDMN. Given that the loadings were in the same direction, although IFLTPR was the highest loading variable in the Checklist factor and the lowest loading variable in the RAMOS factor, a small positive correlation would be expected. As seen in Table 49, positive coefficients are found in each instructional year (averaging about .35).

Second, the Checklist factor for the amount of seatwork (SWT) was based on two variables, ROLEMN and ATINPR, both with equal loadings. These same two variables loaded on the RAMOS direct group instruction (DGI) factor, though two other variables were also components of this factor. Since the loading for ROLEMN was relatively low within the RAMOS factor and the loadings were in the opposite direction, a small negative correlation would be expected. From Table 49, the correlations within the first three instructional years are indeed small, and in a negative direction; the correlation found at the last instructional is positive and large (.75).

Table 49

Instruction - Spanish:
Correlations between the Checklist
and RAMOS Factor Scores for the Bilingual Sample

Instructional Year 0
(N=51)
Checklist

		ADC	STW	PMT	DTC	NST
	QFL	-0.152	-0.063	0.396	-0.248	0.172
	DGI	0.337	-0.193	0.418	-0.227	0.045
	ETT	0.098	-0.072	-0.119	-0.451	0.069
RAMOS	NST	-0.186	0.007	-0.187	0.425	0.356
	ADC	0.118	0.191	0.119	0.182	-0.153
	SMT	0.336	0.173	-0.178	-0.438	-0.058
	CNT	-0.057	0.169	-0.136	-0.404	0.322

Instructional Year 1
(N=64)
Checklist

		ADC	STW	PMT	DTC	NST
	QFL	-0.134	0.382	-0.080	-0.350	-0.314
	DGI	0.101	-0.177	0.298	-0.001	0.527
	ETT	-0.023	0.189	0.042	0.518	0.081
RAMOS	NST	0.222	0.191	-0.089	0.087	0.701
	ADC	0.352	0.519	0.402	-0.246	0.319
	SMT	-0.088	0.085	-0.023	0.345	0.019
	CNT	-0.212	-0.234	-0.243	-0.248	-0.259

Instructional Year 2
(N=60)
Checklist

		ADC	STW	PMT	DTC	NST
	QFL	0.180	0.230	0.180	-0.193	-0.035
	DGI	0.378	-0.052	0.224	0.187	-0.429
	ETT	-0.058	-0.060	0.268	-0.331	0.179
RAMOS	NST	-0.482	0.024	-0.370	0.363	0.221
	ADC	0.181	-0.161	0.182	-0.091	-0.094
	SMT	-0.104	-0.062	-0.128	0.292	0.349
	CNT	-0.205	0.026	-0.122	0.061	0.573

Instructional Year 3
(N=18)
Checklist

	ADC	STW	PMT	DTC	NST
QFL	0.323	0.358	0.463	0.151	-0.276
DGI	-0.112	-0.293	0.000	-0.223	0.181
ETT	-0.292	0.006	-0.129	0.075	-0.396
NST	-0.373	-0.350	-0.223	-0.039	0.628
ADC	0.582	0.328	0.728	0.637	-0.199
SMT	0.724	0.717	0.746	0.666	0.092
CNT	0.071	-0.192	0.426	0.483	0.592

Instructional Year 4
(N=9)
Checklist

	ADC	STW	PMT	DTC	NST
QFL	0.541	-0.276	0.430	-0.242	-0.348
DGI	-0.987	0.754	0.310	0.147	-0.678
ETT	0.672	-0.359	0.370	-0.210	-0.191
NST	0.035	-0.210	-0.703	0.125	0.801
ADC	0.336	-0.097	0.557	-0.211	-0.527
SMT	0.576	-0.273	0.445	-0.204	-0.307
CNT	0.990	-0.748	-0.296	-0.142	0.666

684

Third, two of the three variables constituting the Checklist factor of primary materials (PMT) loaded on the quality of formal language (QFL) factor in the RAMOS set, namely MAT1MN and ATLRMN. Although the latter factor contained three additional variables and the proportional weightings differed between the two data sets, a small positive correlation between factors would be expected. Again from Table 49, the correlation at the first instructional year is zero, becoming positive and increasing in magnitude over the remaining instructional years (from .18 to .45).

Finally, the number of students (NST) factor from the Checklist data consisted of a single loading variable (NSTDMN). This variable, along with two others (ROLEMN and PRDCMN) loaded on the number of students factor from the RAMOS data. Given the same loading direction, a positive correlation would be expected between these factors, and Table 49 shows such positive relationships at each instructional year (about .6).

In summary, again acknowledging the weakness of the analysis, the correlation pattern between factors with shared components gives some support to a common instructional picture being provided by the two instruments. Having described the observed and planned instructional indices derived from the RAMOS and Checklist data sets, the next section describes the derivation and breakdown of a global index of the reading program received which is used in the subsequent integrative analyses.

INSTRUCTIONAL PROGRAM

In the analysis of the relationship between reading growth and instruction (Volume 7), an approximate index of the number of years of Spanish reading received by each target student was employed in conjunction with the dimensions of observed and planned instruction already discussed in this volume. In this section, the method employed to derive this nominal index, and a discussion of the distribution of the number of years of Spanish reading received within each site, are presented.

Program Determination

The first step in deriving the nominal number of years of Spanish reading involved determining the percentage of time devoted to English versus Spanish reading instruction within each semester based on the RAMOS and Checklist data. These individual semester results were then combined to obtain a longitudinal picture of English reading versus Spanish reading instruction. Based on these, the data for each individual student were examined, and the number of consecutive years of Spanish reading instruction was tabulated. Note that any indication that some Spanish reading instruction was received during a given academic year was treated as if the entire year contained instruction in Spanish reading. As such, the resulting program variable is best described as an index of the number of years with some direct exposure to Spanish reading.

Program Analysis

Table 50 presents the distribution of the nominal number of years of Spanish reading instruction received by the target students within each of the study's five bilingual sites. First, note the percentage of students who did not receive reading instruction in Spanish: about a third of the students at Sites 0 and 5, half at Sites 1 and 2, and four-fifths at Site 3. Further, with the exception of Site 3, which shows rapid exit from bilingual reading programs, note that for students who did receive Spanish reading instruction, most stayed in the program for at least two years, more so at Sites 0 and 1 (over 90%) than at Sites 3 and 5 (about 60%).

Finally, note that the numbers of students receiving Spanish instruction under this analysis generally overestimates the numbers indicated in the RAMOS and Checklist data. This simply reflects missing data: for some students within some semesters, only RAMOS data, or only Checklist data were available, and in some cases neither were available (here, decisions about whether the particular student received any Spanish reading instruction were based on data from the surrounding semesters, plus any other available information, including tht concerning what fellow classmates were receiving at the time). Thus, as will be seen in the next volume, the program variable allows a partial assessment of the influence of Spanish reading instruction that complements the assessment made for those students where observed and/or planned instruction data were available.

SUMMARY

This section provides a brief summary of the main findings concerning the dimensions of instruction assessed in this study. The summary will follow the general outline of the volume, first treating the RAMOS indices, then the Checklist indices, and finally the relationships between the two. The section will conclude with a summary of the nominal instructional program variable.

Reading and Mathematics Observation System

The primary source of instructional data for the study was obtained from regular observations of reading periods in classrooms with primary responsibility for providing reading instruction to each of the target students. In the first sections of this volume, a detailed discussion of the classroom observation instrument, the Reading and Mathematics Observation System (RAMOS), was given. That discussion covered the structure of the interview instrument, the data collection and processing procedures, the derivation of summary indices, descriptive statistics on the aggregated measures for the

Table 50

Reading Program:
Distribution of Years of Spanish Reading for Each Site

Spanish Reading Grade Sequence	Site 0	1	2	3	5	Total
None	18	8	16	61	23	126
K only	3	0	5	9	18	35
K-1	0	1	3	1	0	5
K-2	4	5	2	5	29	45
K-3	13	3	8	-	-	24
K-4	17	0	2	-	-	19
Total	55	17	36	76	70	254

687

bilingual sample, descriptions of the procedures followed in deriving factor scores, and the descriptive statistics and inter-correlations associated with the computed factor scores.

As discussed, the RAMOS is a real-time system providing information on the organization of the classroom, the instructor, the content of instruction, and the response of the students to instruction. Such dimensions of instruction have been captured with regard to both their quality and quantity, and for individual target students. From the aggregate analysis of the first order summary indices of these dimensions, the instruction given the bilingual sample in English reading classrooms from first through fourth grade can be summarized as follows:

> Group sizes tended to be slightly smaller in the early grades when compared to the latter two grade levels (13 versus 15 students).
>
> For about two thirds of the instruction observed, some instructor was associated with the target students, generally a teacher operating in a role of facilitation. Although the subject generally being taught was reading, language arts instruction was also observed during these scheduled reading periods.
>
> Only about 10% of the instruction observed constituted time without an instructional focus (e.g., wait time). In the initial two grades, half of the focused instruction was on decoding work, falling to about 30% by the fourth year. In each year, decoding instruction was largely non-explicit (e.g., favoring a focus on whole word recognition over letter-sound recognition). The amount of time spent on word meanings was quite small during the first two years (around 5%), increasing only to about 20% by the fourth year. The instruction that was offered in this area, however, tended to be quite explicit (e.g., favoring a focus on vocabulary enrichment over one on dictionary usage). About 30% of the instructional focus time was on instruction in the meanings of sentences and texts, with only a slight increase over the years. This instruction was generally non-explicit within each year (e.g., favoring a focus on literal facts over making inferences).
>
> Primary materials were observed in use about 90% of the time in each year, with an increase toward more text usage. Ancillary materials were employed about 60% of the time during the initial two instructional years, increasing to about 70% in the final two years. The quality index of these materials tended to be lower than the primary materials (e.g., tending more toward paper and pencil than text).
>
> About 90% of the observed time contained instructional activity/tasks. Independent work accounted for about half the instructional time during the first two years, dropping to about 35% in the following two years. Conversely, group work

represented about 40% of the instructional time during the first two years, increasing to about 50% for the last two years. For independent work, the level of formal language demand increased over the instructional years, starting at a relatively low level in the first year, and increasing to mid-level by the last two years. For group work, however, little change in level of formal language demand occurred; it was, however, noticeably higher than that associated with independent work.

For the student response measures, the number of nonengaged students was low, as was the rated noise level, while productivity was rated as medium in each year.

Keeping in mind the cautions concerning the interpretation of the latter years of the Spanish reading instruction data (given the greatly reduced sample size), the aggregate descriptive data for the bilingual sample can be summarized as follows:

The number of students associated with the observed instructional groups tended to be slightly smaller in the early years (around 13 students, as in the English data) but much larger in later years (around 20 students).

During the first year, some instructor was associated with the target students for about 60% of the instruction observed, increasing to over 80% in the following years. This instructor, as in the English observations, generally was a teacher, and the role played tended to be one of facilitation. The subject being taught generally required substantial reading, but as in the English data, the teaching of other subjects was also observed.

About 90% of the observed instruction contained instructional foci. About half of this focused time was devoted to decoding instruction during the first and second instructional years, falling to about 30% by the fourth year, mirroring the English observations. This decoding instruction also tended to be non-explicit with respect to letter-sound correspondence specificity. The amount of time spent on word meanings was negligible during the first year (only 2%), increasing to about 10% in the second year, and to 20% in the last two years; the limited amount of instruction that was offered in this area, however, tended to be quite explicit (as it was in the English data). For each year, about 30% of the instructional focus time was devoted to instruction in the meanings of sentences and texts; the quality of this instruction was generally non-explicit.

For materials, the use and quality of primary materials closely resembled that found in the English reading observations with such materials employed about 90% of the time and a trend toward more textual material usage over years. Ancillary materials were used about 60% of the time during the first two instructional years, and about 75% for years 3 and 4. As in the English observations, the quality index of these materials was lower than that associated with the primary materials.

For the activity/task scales, about 90% of the observed time contained instructional activity/tasks. Independent work accounted for about half the instructional time during the first year, dropping to about 35% in the following two years, and to zero in the final year. Conversely, group work represented about 40% of the instructional time during the first year, increasing to about 60% for the next two ye - and to 85% for the final year. For independent work, the -. of formal language demand started at a relatively low level in the first year, and increased to mid-level by the last year. For group work, however, little change in level of formal language demand occurred; it was, as in the English data, noticeably higher than that associated with independent work.

For the student response measures, the number of nonengaged students was generally low, but substantially higher in the latter two years; productivity was rated as medium in each year, while noise tended to be low.

In general, the descriptions of reading instruction in English programs versus those in Spanish programs based on these aggregate data did not substantially differ. In both, instruction was largely conducted by a teacher acting in a role of facilitation. Over instructional years, an increased reliance on group work over independent work was seen. Much of the early grade work was focused on decoding, declining in the latter instructional years. The quality of this instruction tended to be non-explicit as little instruction dealing explicitly with letter-sound correspondences was seen. Little instruction in vocabulary was observed, although that which was observed tended to be quite explicit. Finally, instruction in sentence/text meaning complemented the time devoted to decoding, showing a small increase over instructional years; like decoding, this instruction was generally non-explicit, favoring a focus on literal facts.

The summary data which these descriptions were based upon were subjected to factor analyses in order to reduce the number of instructional indices. In both the English and Spanish data sets, seven factors were derived. The seven factors identified in the English data analysis included the following:

(1) Engaged Text Time, an index of reading time where students were engaged with text materials,

(2) Direct Group Instruction, an index of direct instruction delivered by an instructor which was aimed at groups of students, rather than individuals,

(3) Quality of Formal Language, a measure of the formal language demands made upon the students,

(4) Amount of Decoding, a measure of the relative amount of time devoted to instruction in decoding,

(5) Productivity, an index of the conditions promoting high individual student productivity,

(6) Secondary Materials, a measure of relative usage of secondary materials (though this interpretation is oversimplified), and

(7) Number of Students, an index of the number of students constituting an instructional group.

The factor solution derived from the Spanish data was similar to that derived from the English data, with five of the seven factors containing many of the same component variables. The seven factors identified in the Spanish factor analysis included:

(1) Quality of Formal Language (corresponding to the third English factor derived),

(2) Direct Group Instruction (the second English factor),

(3) Engaged Text Time (the first English factor),

(4) Number of Students (the last English factor),

(5) Amount of Decoding (the fourth English factor),

(6) Secondary Materials, an index of both the quality and quantity of secondary material usage (only tangentially related to the sixth English factor), and

(7) Control, a complex factor without an English correspondence, which is essentially an index of the number of management interruptions.

Based on these noticeably similar factor solutions, the procedures followed in the computation of factor scores were discussed. Such scores constitute the major set of observed instruction indices employed in the integration analysis that is the subject of the next volume.

Teacher Checklists

A secondary source of instructional data was obtained from regular interviews with the teachers who were primarily responsible for providing reading instruction to each of the target students. Earlier sections provided detailed discussions of the instrument, the Teacher Checklists (Checklists), including the structure of the interview instrument, the data collection and processing procedures, the derivation of summary indices, descriptive statistics on the aggregated measures for the bilingual sample, descriptions of the procedures employed in the derivation of factor scores, and the descriptive statistics and inter-correlations associated with these later measures.

The Checklist was modeled after the RAMOS, using its coding scheme to capture the teacher's instructional plans for individual target students. Again, the relevant dimensions of instruction concerned the organization of the classroom, the instructor, and the content of instruction. From the aggregate analysis of the first order summary indices of these dimensions, the instruction given the bilingual sample in English reading classrooms from the first through fourth grade can be summarized as follows:

From the aggregate analysis of the first order summary indices, planned instruction for the bilingual sample in English reading classrooms can be briefly summarized as follows:

> About three-fourths of the instructional time planned was devoted to decoding during the first and second instructional years, falling by about 15% in <u>each</u> of the subsequent two years; with respect to letter-sound pairings, this planned instruction was largely non-explicit in each of the grades. The amount of time to be spent on word meanings was around 5% in first grade, increasing by about 5% with each subsequent year. The instruction that was planned was quite explicit in the first year, with decreasing explicitness in each of the subsequent years. About 20% of the planned instructional focus was devoted to instruction in the meanings of sentences and texts for the first two years, increasing by about 8% in each of the two subsequent years. The quality of this planned instruction was generally non-explicit (e.g., favoring a focus on literal facts over making inferences).
>
> For materials, there was a clear trend toward more planned text usage in the latter years.
>
> About 30% of the planned instructional activities consisted of independent work in each of the instructional years. Conversely, group work represented about 70% of the planned activities. For independent work, the level of formal language demand was, on average, mid-level; for group work, the level of formal language demand was noticeably higher.
>
> In each of the four instructional years, the instructor was primarily responsible for the delivery of the planned instruction, and the role to be played tended to be mid-way between one of facilitation and direct instruction.
>
> There appeared to be a trend toward larger groups over the instructional years, with an average size of about 9 students in the initial years, and about 13 students in the later years.

The average amount of planned daily reading instruction was about 40 minutes in the initial year, increasing to about 60 minutes in the later years.

From a similar aggregate analysis, the planned instruction for the bilingual sample participating in Spanish reading classrooms can be summarized as follows, again cautioning the interpretation of the data from the last two years:

About three-quarters of the instructional time planned was devoted to decoding instruction during the first and second years, matching the planned English instruction. However, whereas this value fell to about 45% in the English data, it dropped only to about 60% in the Spanish data. The quality of this Spanish decoding instruction matched that of the English instruction, namely, a non-explicit focus with respect to letter-sound pairings. The amount of time to be spent on word meanings was about 5% in the first and second instructional years, and about 15% in the third and fourth years, showing a different pattern than that found in English. However, the instruction planned was relatively explicit in each year, not showing the steady decline found in the English data. About 15% of the planned instructional focus was devoted to instruction in the meanings of sentences and texts for the first three years, increasing by about 10% in the fourth year -- for these students, the relatively larger amounts of time planned for decoding instruction were associated with relatively smaller amounts of planned time for the study of connected text. The explicitness of this instruction tended to be somewhat lower than that found in English.

For materials, as in English, there was a clear trend toward more planned text usage over instructional years.

About 25% of the planned instructional activities consisted of independent work during the first two years, dropping to about 15% in the remaining two years. Conversely, group work represented about 75% of the planned instruction during the first two years, increasing to about 85% in instructional years three and four. The relative level of formal language demand associated with this work mirrored that found in the English data: independent work about mid-level with group work higher in formal demands.

As in the English data, a teacher was primarily responsible for the delivery of the planned instruction The planned role of the instructor tended to be mid-way between facilitation and direct instruction during the first two years; however, in the subsequent years this role tended to be substantially higher (reflecting almost exclusive direct instruction).

As was found in the English data, the number of students expected in each group tended toward larger groups in latter years (representing an increase from about 9 to 13 students).

The average amount of planned daily reading instruction was about 30 minutes in each instructional year except the second, where it was almost 40 minutes per day. Thus, the amount of planned reading time for students in Spanish reading programs was about half that planned for students in English reading programs.

Based on these aggregate data, the descriptions of reading instruction in English programs versus Spanish programs are very similar. In both, instruction was largely to be conducted by a teacher acting in a role mid-way between facilitation and direct instruction. Over instructional years, group work was relied upon more than independent work. Much of the early grade work was to be focused on decoding, it's emphasis declining in the latter instructional years. The quality of this instruction tended to be non-explicit as little instruction dealing explicitly with letter-sound correspondences was planned. Little instruction in vocabulary was planned, although that which was planned tended to be quite explicit. Finally, instruction in sentence/text meaning complemented the time devoted to decoding, showing small increases over instructional years; like decoding, this instruction was generally non-explicit, favoring a focus on literal facts.

In spite of these similarities, there were some important differences between the planned English and Spanish instruction. First, the amount of planned decoding time in Spanish did not decline as precipitously as that for English, suggesting that the students enrolled in Spanish programs were seen by their teachers as being more deficient in decoding skills, and thus requiring more instruction. Complementing this difference, less time in the Spanish reading programs was to be devoted to instruction in sentence/text meaning. Finally, the total amount of planned reading instruction was about half that planned for English reading.

The summary indices derived from the Checklist data were subjected to factor analyses in order to reduce the number of instructional indices. In both the English and Spanish data sets five factors were derived, and those identified in the English data analysis included the following:

(1) Amount of Comprehension, a measure of the relative amount of planned time to be devoted to instruction in comprehension,

(2) Quality of Formal Language, an index of the formal language demands required by the planned instruction,

(3) Seatwork, an index of the relative amount of time to be devoted to independent seatwork as opposed to group work,

(4) Primary Materials, an index of the planned usage of primary materials (although this factor is more complex), and

(5) Group Vocabulary, an index of the relative amount of time to be devoted to group instruction in the meaning of words.

in the Spanish data set, the five factors derived included:

(1) Amount of Decoding (a compliment of the first English factor),

(2) Seatwork (corresponding to the third English factor),

(3) Primary Materials, an index of the quality of the primary materials to be employed (although this factor is more complex),

(4) Decoding Teacher Classification, a factor defining the relative educational training of the teacher expected to deliver decoding instruction (which was associated with the explicitness of such planned instruction), and

(5) Number of Students, an index of the relative instructional group size.

Again, based on these analyses, factor scores were computed for each student within each instructional year, and these constitute the major set of planned instruction indices employed in the integration analysis that is the subject of the next volume.

Relationships between RAMOS and Checklist Indices

Although not of primary interest in this study, the relationship between observed instruction, as quantified in the RAMOS summary measures, was contrasted with planned instruction, as quantified in the Checklist summary measures. For the English data, the aggregated contrasts over the final four instructional years, based on the common summary indices represented in the two instruments can be summarized as follows:

For the number of students, the trend toward increasing group size was found under both assessments, though the Checklist data tended to underestimate what was actually observed.

Considering the instructor variables, both assessments found the teacher with primary responsibility for the delivery of instruction, though the role played tended to be overestimated in the Checklist data.

For the instructional focus variables, the same trend was found in both assessments with respect to the quality indices (concerned instruction in decoding, word meaning, and sentence/text meaning). However, slightly more decoding instruction was planned than observed, and, concomitantly, slightly less instruction on connected text was planned than observed.

For the activity/task scales, the same trend was generally found in both assessments with respect to the quality indices (concerned with the formal language demands of independent work and of group work); however, more independent work (and thus, less group work) was observed than planned.

For the materials, the trend toward more usage of text was found in both instruments.

In general these aggregate contrasts suggest that the instructional picture provided by the RAMOS data is not substantially different from that provided by the Checklist data. The most noticeable discrepancies concern the estimate of the instructor's role and the relative proportions of time devoted to particular instructional foci and activity/tasks; the most noticeable similarities are those related to the quality estimates of instruction (in particular, for the instructional foci and activity/tasks).

The Spanish aggregate contrasts between the RAMOS and Checklist indices over the final four instructional years based on the common summary indices represented in the two instruments can be summarized as follows:

For the number of students, the trend toward increasing group size was found under both assessments, though the Checklist data, as was found in the English data, tended to underestimate the average group size actually observed.

Considering the instructor variables, both assessments found the teacher with primary responsibility. Again, as in the English data, the role played tended to be overestimated in the Checklist data, substantially so with respect to the last two years of instruction.

For the instructional focus variables, the same trend was found in both assessments with respect to the quality indices. However, more decoding instruction was planned than observed (substantially more in the latter two years), and, concomitantly, less instruction on connected text was planned than observed.

For the activity/task scales, again, the same trend was generally found in both assessments with respect to the quality indices; however, more independent work (and thus, less group work) was observed than planned.

For the materials, the trend toward more usage of text with increasing instructional years was found in both instruments (as was found in the English data).

Thus, as in the English data, the aggregate contrasts suggest that the instructional pictures provided by the two data sets are similar, again, with greater discrepancies found in the quantity indices than in the quality indices. A better picture of the correspondence between the data sets (though not the best assessment) is to be found in the pattern of factor score inter-correlations, which, as discussed earlier, supports their convergence.

Instructional Program

In the last section of the volume, the procedures followed in deriving a student index concerned with the number of years with some direct exposure to Spanish reading were discussed The analysis of the resulting measure suggested the following. First, a large percentage of students did not receive any reading instruction in Spanish, although each was enrolled in a bilingual program (though not necessarily one with a Spanish reading component) when initially selected for study participation. Further, with the exception of one site which showed rapid exit from bilingual reading programs, students who did receive Spanish reading instruction were likely to stay in those reading programs for at least two years. Finally, the discrepancies between the number of cases identified through this measure as opposed to the RAMOS and Checklist indices were discussed.

This concludes the presentation of the instruction data. Having earlier discussed the language data (Volume 4) and the reading data (Volume 5), the next volume will treat their integration.

REFERENCES

Calfee, R. C. & Calfee, K. H. (1974). Reading and mathematics observation system - RAMOS. Unpublished manuscript, Stanford University: Stanford, CA.

Calfee, R. C., & Calfee, K. H. (1976). Reading and mathematics observation system - RAMOS/II (rev.). Unpublished manuscript, Stanford University: Stanford, CA.

Calfee, R. C., & Calfee, K. H. (1978). Reading and mathematics observation system - RAMOS/SEDL. Unpublished manuscript, Southwest Educational Development Laboratory: Austin, TX.

Piontkowski, D. C. (1981). A structural analysis of classroom processes in beginning reading instruction. Unpublished manuscript, Stanford University: Stanford, CA.

APPENDIX A

Bilingual Classroom Questionnaire
and
Inventory of Bilingual Instruction

BILINGUAL CLASSROOM QUESTIONNAIRE

Southwest Educational Development Laboratory
Division of Bilingual and International Education
211 East Seventh Street
Austin, Texas 78701

The responses from this questionnaire will be used to design inservice education for teachers, and will not be used to evaluate teacher knowledge, skills, or attitudes.

January 1979

BEST COPY AVAILABLE

Teacher Name ______________________

School ______________________

Grade Level(s) ______________________

Date Completed ______________________

INSTRUCTIONS

The Bilingual Classroom Questionnaire will be used to describe instructional practices in bilingual classrooms. The Questionnaire is part of a project designed (1) to provide educators with procedures for describing the type of bilingual education in their schools and (2) to identify staff development needs for applying bilingual education successfully.

The Questionnaire will take about 30 minutes to complete. Information is requested in six columns. The example provided below shows how to complete columns one through three. The discussion which follows describes procedures for completing columns one through three and adds information on columns four through six.

In Column 1, Current Daily Schedule, please list in time sequence the daily activities of the students in your classroom. If more than one activity occurs during a given time period, list each of the concurrent activities separately. For example, suppose that from 8:00 to 8:30 one group of students receives Spanish reading instruction while another group receives English oral language development. Each activity would be listed separately, as shown in the example.

1		2		3					
CURRENT DAILY SCHEDULE (for your homeroom students)		ANTICIPATED DURATION		LANGUAGE CATEGORIES (of Students within Instructional Groups)					
HOURS	ACTIVITY	Continuous (All Year Long)	Noncontinuous (Please indicate—e.g., 1 day each week; every other week, 2 weeks out of every 4, etc.)	BB	BE	BS	ME	MS	LL
8:00-8:30	Spanish Reading	✓				✓		✓	
	English Oral Language Development	✓		✓	✓				
8:30-9:00	Science		2 weeks out of every 4	✓		✓			
8:30-9:00	Social Studies		2 weeks out of every 4	✓		✓			
9:00-9:30	P.E.		4 days each week	✓	✓	✓	✓	✓	
9:00-9:30	Art		1 day each week		✓	✓	✓	✓	

In Column 2, Anticipated Duration, indicate whether the scheduled activity occurs throughout the year (✓) or on a more limited basis (e.g., one day each week, two weeks out of every four, etc.). If different activities are scheduled during the same period but on a rotating basis, please list all the activities as shown in the example above (e.g., from 8:30-900 Science is taught for two weeks with Social Studies being taught the following two weeks before the cycle repeats itself; from 9:00-9:30 P.E. is taught for four days each week while Art is taught on the remaining day).

The Language Categories noted in Column 3 are to be completed for every Activity noted in Column 1. For each Activity, check the Language Categories of the students participating in the activity. The Language Category Definitions and abbreviations are listed on the following page.

701

Language Category Definitions

Balanced Bilingual (BB) — Totally fluent in both English and Spanish.

Partial Bilingual, English Dominant (BE) — Understands all spoken English and produces English utterances with native-like fluency and correctness in syntax (grammar) and vocabulary. Also understands some spoken Spanish and can produce fairly complete sentences in Spanish but with less than native-like fluency. His/her sentences in Spanish are somewhat awkward with regularized errors in syntax and vocabulary.

Partial Bilingual, Spanish Dominant (BS) — Understands all spoken Spanish and produces Spanish utterances with native-like fluency and correctness in syntax (grammar) and vocabulary. Also understands some spoken English and can produce fairly complete sentences in English but with less than native-like fluency. His/her sentences in English are somewhat awkward with regularized errors in syntax and vocabulary.

Monolingual English (ME) — Understands all spoken English and speaks English with ease and complete native-like fluency and correctness. If any Spanish is understood or spoken it is no more than a few isolated words or expressions.

Monolingual Spanish (MS) — Understands all spoken Spanish and speaks Spanish with ease and complete native-like fluency and correctness. If any English is understood or spoken it is no more than a few isolated words or expressions.

Limited English/Limited Spanish (LL) — Does not have native competence in either English or Spanish. It may appear that he/she understands spoken English and Spanish but the oral production in both languages is labored, characterized by awkward sentences and systematic errors in syntax (grammar) and vocabulary.

In Column 4, the Primary Instructor of the Instructional Activity should be indicated. Alternatives are the Teacher, Team Teacher, Resource Teacher, Teacher Aide, and Other. Select one (✓) of these per Activity noted in Column 1.

In Column 5, please check the Language of Instruction for each Activity listed in Column 1. Definitions of the four alternatives are listed below. Select a single category for each Instructional Activity.

Language of Instruction

Primarily Spanish. Instruction is provided exclusively in Spanish or primarily in Spanish with only an occasional use of English during the instructional period.

Primarily English. Instruction is provided exclusively in English or primarily in English with only an occasional use of Spanish during the instructional period.

Alternating Use of Both Languages. Both languages are used approximately an equal amount of time during the instructional period. As distinguished from code-switching, alternating use of the two languages is characterized by exclusive use of one language at a time during an instructional event.

Code-Switching. This form of language involves introducing into the context of one language stretches of speech that exhibit the other language's phonological and morphological features.

In Column 6, indicate the Language of Materials for each Instructional Activity. The alternatives are English, Spanish, Both, or No Material. Select one (✓) of these for every Activity noted in Column 1.

*KEY: BB = Balanced Bilingual
BE = Partial Bilingual, English Dominant
BS = Partial Bilingual, Spanish Dominant

ME = Monolingual English
MS = Monolingual Spanish
LL = Limited English/Limited Spanish

1		2		3							
CURRENT DAILY SCHEDULE (Of your homeroom students)		ANTICIPATED DURATION		LANGUAGE CATEGORIES* OF STUDENTS WITHIN INSTRUCTIONAL ACTIVITIES						PRIMARY I INSTRU (Che	
		Continuous (All Year Long)	Noncontinuous (Please indicate, e.g., 1 day each wk., every other wk., 2 wks. out of every 4, etc.)								
HOURS	ACTIVITY			BB	BE	BS	ME	MS	LL	Teacher	Team Teach

BEST COPY AVAILABLE

703

QUESTIONNAIRE

			5				6			
CTOR OF ACTIVITY (e)			LANGUAGE OF INSTRUCTION (Check One)				LANGUAGE OF MATERIALS (Check One)			
source acher	Teacher Aide	Other	Primarily English	Primarily Spanish	Alternating Use of Both	Code Switching	English	Spanish	Both	No Material

BEST COPY AVAILABLE

Instructions to the Data Collector

INVENTORY OF BIL L INSTRUCTION

Southwest Educational Development Laboratory
Division of Bilingual and International Education
211 East Seventh Street
Austin, Texas 78701

April, 1981

CONFIDENTIAL: The responses to this inventory will be treated confidentially

INVENTORY OF BILINGUAL INSTRUCTION

GENERAL DIRECTIONS: The Inventory of Bilingual Instruction (IBI) is a descriptive survey instrument developed to record instructional practices in bilingual environments. Data from PART ONE of the IBI will describe typical curriculum patterns. PART TWO of the IBI will be used to record the instructor's beliefs about why children learn.

The interview is to be conducted by a data collector who will interview one teacher at a time. The interview will take about 60 minutes. Before the interview each data collector will check equipment (tape recording is preferred, but, maybe optional), have pencil and instrument(s) on hand, and arrange for the interview to be conducted at the site. The instructor's name, school, etc. (see cover page of the IBI) will be filled out prior to starting the interview.

The data collector will ask the questions that follow and record the information directly on the instrument. EACH question is identified by letter, i.e., Question A, B,...G. for part one of the interview. Exact questions are given. Additional questions (probes) are included to clarify or complete each item. Probes are identified by this symbol ( ). A triangle △ will indicate what the data collector will do to record the data. Questions and procedures to complete part two of the IBI are included separately.

NOTE: DEFINITIONS OF LANGUAGE CLASSIFICATIONS (see column 3, part one) should be discussed with the instructor as necessary. Definitions follow:

Monolingual Spanish* -- Understands all spoken Spanish and speaks Spanish with ease and complete native-like fluency and correctness. If any English is understood or spoken it is no more than a few isolated words or expressions.

Spanish Dominant -- Understands all spoken Spanish and produces Spanish utterances with native-like fluency and correctness in syntax (grammar) and vocabulary. Also understands some spoken English and can produce fairly complete sentences in English but with less than native-like fluency. His/her sentences in English are somewhat awkward with regularized errors in syntax and vocabulary.

Bilingual -- Totally fluent in both English and Spanish.

English Dominant -- Understands all spoken English and produces English utterances with native-like fluency and correctness in syntax (grammar) and vocabulary. Also understands some spoken Spanish and can produce fairly complete sentences in Spanish but with less than native-like fluency. His/her sentences in Spanish are somewhat awkward with regularized errors in syntax and vocabulary.

**Monolingual English -- Understands all spoken English and speaks English with ease and complete native-like fluency and correctness. If any Spanish is understood or spoken it is no more than a few isolated words or expressions.

Other -- The student speaks or understands a language other than English, Cantonese, Navajo, or Spanish.

* The data collector will SUBSTITUTE THE APPROPRIATE LANGUAGE (Cantonese or Navajo) FOR SITES WHERE THESE LANGUAGES ARE SPOKEN IN THE HOMES OF NON-ENGLISH SPEAKERS. Spanish is used here for the sake of brevity.

** Includes monolingual speakers of any ethnic group, e.g. Anglo students and students of the target ethnic group but not speakers of that ethnic group language. Language, not ethnicity, is criteria for classification.

DIRECTIONS FOR PART ONE

Questions are arranged to correspond with the items in the page titled "TEACHER CURRICULUM INTERVIEW."

Each question is identified by letter. Exact questions are given. A symbol (Δ) indicates what the data collector is to do to record the data.

Begin the interview by introducing yourself. Tell the instructor "The purpose of the survey is to describe your curriculum."

COLUMN 1

Δ Ask Question A) What time do you start your class and what is the first subject that you teach? (I'm going to ask you about your curriculum schedule. Please tell me what the first subject you teach is and the time this is done?)

Δ ENTER Response in Column 1 to indicate time subject is taught.

COLUMN 2

Δ ENTER name of subject in Column 2 above the word "Group". ENTER each subject separately.

Δ Ask Question B) Do you teach this subject ____________ to the whole class at the same time or to small groups of students?

...If small groups, how many groups are there? Use space provided for Group 1, 2,...5. Ask next question.

Δ Ask Question C) Do you teach this all semester long? If YES, mark a (✓) under "duration". If NO, write in response given by instructor.

Δ Ask Question D) How many days per week is this taught?

ENTER number of days per week.

COLUMN 3

Δ Ask Question D) Do you have children of different language classification? If Yes, how many of these are Monolingual Spanish; Spanish Dominant; Bilingual, English Dominant; Monolingual English; Other.

Δ READ definition of each from "General Directions" page 1.

IF ONLY ONE LANGUAGE CLASSIFICATION, ENTER NUMBER OF STUDENTS under the appropriate classification.

IF MORE THAN ONE LANGUAGE CLASSIFICATION, ENTER NUMBER OF STUDENTS per language classification in the space provided.

COLUMN 4

Δ Ask Question E) Who provides the instruction for this subject ____________ (who is the primary instructor for the ____________ subject?)

T = teacher	V = volunteer	R = resource
A = aide	P = peer	SE= special education teacher
TT= team teacher	C = combination	

Δ ENTER the appropriate code in column 4.

BEST COPY AVAILABLE

7)7

COLUMN 5

△ Ask Question F) What is the primary language of instruction (primarily English; Primarily Spanish; both?)

△ ENTER response in Column five.

COLUMN 6

△ Ask Question G) Do you use written materials for this instruction, _______________ if so, what language are the materials written in (primarily English; primarily Spanish; both; no materials?)

△ ENTER response in Column six.

*REPEAT QUESTIONS B thru G for next subject. Complete the total instructional day for each teacher interviewed.

708

BEST COPY AVAILABLE

Name ______________________

Grade ______________________

School ______________________

Site ______________________

Date ______________________

INVENTORY OF BILINGUAL INSTRUCTION

(Language & Literacy Learning in Bilingual Instruction)

Division of Bilingual and International Education
211 East Seventh Street
Austin, Texas 78701

July 1981

				# of Students Per Classification							Lang. of Instruction		Lang. of Materials		
Time	Content	Grouping	Days per Week	Home Language Monolingual	Home Language Dominant	Bilingual	English Dom.	Eng Monolingual	Other	Instruction/ Instructor	English	Home Language	English	Home Language	No Materials
		1													
		2													
		3													
		4													
		5													
		1													
		2													
		3													
		4													
		5													
		1													
		2													
		3													
		4													
		5													
		1													
		2													
		3													
		4													
		5													
		1													
		2													
		3													
		4													
		5													

BEST COPY AVAILABLE

710

				# of Students Per Classification							Lang. of Instruction		Lang. of Materials		
Time	Content	Grouping	Days per Week	Home Language Monolingual	Home Language Dominant	Bilingual	English Dom.	Eng Monolingual	Other	Instruction/ Instructor	English	Home Language	English	Home Language	No Materials
		1													
		2													
		3													
		4													
		5													
		1													
		2													
		3													
		4													
		5													
		1													
		2													
		3													
		4													
		5													
		1													
		2													
		3													
		4													
		5													
		1													
		2													
		3													
		4													
		5													

711

APPENDIX B

Survey of Teachers' Background and Language Skills

SURVEY OF TEACHERS' BACKGROUND AND LANGUAGE SKILLS

SOUTHWEST EDUCATIONAL DEVELOPMENT LABORATORY
211 E. 7th Street, Austin, TX 78701

(Label containing Name, School, District)

DATE DUE:

To make sure that your response is treated confidentially, we will remove this label after we have processed your completed form. Thus, no one will be able to identify you, your school, or district with any information on this questionnaire, after the survey has been completed.

1. WHAT WAS YOUR AGE ON SEPTEMBER 1, 1982 ? (Check appropriate category)

(1) ( ) Under 21
(2) ( ) 21 - 30
(3) ( ) 31 - 45
(4) ( ) 46 - 55
(5) ( ) 56 and over

2. WHAT IS YOUR SEX?

(1) ( ) Female
(2) ( ) Male

3. WHAT IS YOUR ORIGIN OR DESCENT? (Check as many as apply)

(01) ( ) German
(02) ( ) Italian
(03) ( ) Irish
(04) ( ) French
(05) ( ) Polish
(06) ( ) Russian
(07) ( ) English
(08) ( ) Scottish
(09) ( ) Welsh

Hispanic — Mexican
(10) ( ) Mexican American — Mexicano
(11) ( ) Puerto Rican — Chicano
(12) ( ) Cuban
(13) ( ) Central or South American
(14) ( ) Other Hispanic
(15) ( ) Portuguese

(16) ( ) American Indian/Alaskan native
(17) ( ) Black

Asian
(18) ( ) Filipino
(19) ( ) Cambodian
(20) ( ) Chinese
(21) ( ) Japanese
(22) ( ) Korean
(23) ( ) Laotian
(24) ( ) Vietnamese
(25) ( ) Pacific Islander
(26) ( ) Scandinavian Group
(27) ( ) Arabic Group
(28) ( ) Greek
(29) ( ) Other Group not listed (Specify)

____

(98) ( ) I don't know

YOUR EDUCATION AND EXPERIENCE

4. WHAT WAS THE HIGHEST DEGREE YOU EARNED PRIOR TO SEPTEMBER 1, 1982 ?

(01) ( ) A doctoral degree (Ph.D., Ed.D., etc.)
(02) ( ) Specialist degree (A post-master's degree or other than doctorate)
(03) ( ) A master's degree or equivalent
(04) ( ) A bachelor's degree
(05) ( ) Less than a bachelor's degree
(06) ( ) Other (Specify) ______________________

5. AT WHAT GRADE LEVEL DID YOU TEACH DURING THE 1982-83 SCHOOL YEAR?

(01) ( ) kindergarten
(02) ( ) first
(03) ( ) second
(04) third
(05) fourth
(06) Other (specify) ____________

6. WHAT SUBJECT AREAS DID YOU TEACH DURING THE 1982-83 SCHOOL YEAR? (Check all that apply.)

(01) ( ) Art
(02) ( ) Language Arts (other than reading)
(03) ( ) Math
(04) ( ) Music
(05) ( ) Physical Education
(06) ( ) Reading (English)
(07) ( ) Reading (Spanish)
(08) ( ) ESL
(09) ( ) SSL
(10) ( ) Science
(11) ( ) Social Studies
(12) ( ) Special Education
(13) ( ) Other (specify) ____________

7. INCLUDING THE 1982-83 SCHOOL YEAR, HOW MANY YEARS HAVE YOU TAUGHT (total teaching experience)?

(01) ( ) Less than 3
(02) ( ) 3 - 5
(03) ( ) 6 - 10
(04) ( ) 11 - 20
(05) ( ) More than 20

8. INCLUDING THE 1982-83 SCHOOL YEAR, HOW MANY YEARS HAVE YOU TAUGHT IN YOUR PRESENT SCHOOL?

(01) ( ) Less than 3
(02) ([illegible]
(03) [illegible] 0
(05) ( ) 11 - 20
(06) ( ) More than 20

9. INCLUDING THE 1982-83 SCHOOL YEAR, HOW MANY YEARS HAVE YOU TAUGHT IN A BILINGUAL EDUCATION PROGRAM?

(01) ( ) Never
(02) ( ) Less than 3
(03) ( ) 3 - 5
(04) ( ) 6 - 10
(05) ( ) More than 10

10. WHAT IS YOUR TEXAS EDUCATION AGENCY CERTIFICATION STATUS? (Check one)

(01) ( ) State Certified Teacher with Bilingual Endorsement.
(02) ( ) State Certified Teacher with Special Assignment Permit.
(03) ( ) State Certified Teacher with No Bilingual Endorsement or Special Assignment Permit
(04) ( ) Currently teaching on an Emergency Certificate
(05) ( ) Other (specify) ____________

LANGUAGE BACKGROUND AND ABILITIES

11. WHAT LANGUAGE WAS USUALLY SPOKEN IN YOUR HOME WHEN YOU WERE A CHILD? (Check one)

    (01) ( ) English primarily
    (02) ( ) Spanish primarily
    (03) ( ) Spanish and English about equally
    (04) ( ) Other (specify) ____________________

12. COULD YOU SPEAK ANY LANGUAGE(S) OTHER THAN ENGLISH PRIOR TO SEPTEMBER 1, 1981?

    (01) ( ) Yes — Continue with Question 13

    (02) ( ) No — Skip to Question 14, Relevant College or Inservice Training

13. IN THE CHART BELOW, CHECK ALL THE LANGUAGES LISTED IN COLUMN 1 THAT YOU SPOKE PRIOR TO SEPTEMBER 1, 1982. THEN ON THE RIGHT SIDE OF THE CHART, PLACE A CHECK BELOW THE STATEMENT OR STATEMENTS WHICH DESCRIBE THE SITUATION OR SITUATIONS IN WHICH YOU LEARNED THE LANGUAGE(S) YOU CHECKED.

LANGUAGE	SITUATION IN WHICH LEARNED					
	It was usually spoken in my home when I was a child	I acquired it (informally) while living in a country where it is spoken	I acquired it (informally) in a community in the U.S.	I studied it in school in a country where it was spoken	I studied it in school as a foreign language	Other (Write in next to language)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
(01) ( ) English						
(02) ( ) Spanish						
(03) ( ) Other (specify)						

14. FOLLOWING ARE SHORT, DESCRIPTIVE LANGUAGE SITUATIONS. PLEASE RESPOND TO EACH OF THEM ON THE BASIS OF YOUR OWN LANGUAGE ABILITIES.

    A. Imagine that while shopping at a local grocery store you meet your elderly neighbor who speaks primarily Spanish. She addresses you in Spanish.
    MY VERBAL SKILLS ARE SUCH THAT: (Please check one)

    (01) ( ) I would be completely unable to carry on an informal conversation with her in Spanish.
    (02) ( ) I would be able to carry on a conversation with her with limited facility.
    (03) ( ) I would have only minor difficulties in carrying on a conversation with her.
    (04) ( ) I would be able to carry on a conversation with her with complete confidence and ease.

B. Imagine that you are to have a meeting with a parent who has recently arrived from Mexico. The parent speaks no English. The purpose of the meeting is to describe to the parent the type of curriculum which her son will have in your classroom.

MY VERBAL SKILLS IN SPANISH ARE SUCH THAT: (Please check one)

(01) ( ) I would be unable to carry out the above task.
(02) ( ) I would be able to carry out the above task with limited facility.
(03) ( ) I would ha:e only minor difficulties in carrying out the above task.
(04) ( ) I would be able to carry out the above task with complete confidence and ease.

B. Imagine that you were asked to make a presentation in Spanish to your local or state professional organization in which you describe successful teaching activities that you use in your classroom.

MY VERBAL SKILLS IN SPANISH ARE SUCH THAT: (Please check one)

(01) ( ) I would be unable to make such a presentation in Spanish.
(02) ( ) I would be able to make such a presentation in Spanish with only limited facility.
(03) ( ) I would be able to make such a presentation in Spanish with only minor difficulties.
(04) ( ) I would be able to make such a presentation in Spanish with complete confidence and ease.

D. A close friend of yours has just received a letter from a monolingual Spanish-speaking cousin who lives in Mexico and who your friend sees infrequently. While your friend speaks Spanish fluently, she is somewhat insecure in her ability to read Spanish. She has asked you to read the letter and to verify for her the content of the letter.

MY READING SKILLS IN SPANISH ARE SUCH THAT: (Please check one)

(01) ( ) I would be unable to read the letter.
(02) ( ) I would be able to read the letter with limited facility.
(03) ( ) I would be able to read the letter with only minor difficulties.
(04) ( ) I would be able to read the letter with complete confidence and ease.

E. You have been asked by your school librarian to preview some Spanish language newspapers and current event magazines and to advise her as to the quality of the publications and their appropriateness for circulation in a public school library.

MY READING SKILLS IN SPANISH ARE SUCH THAT: (Please check one)

(01) ( ) I would be unable to read any of the publications.
(02) ( ) I would be able to read the publications with limited facility.
(03) ( ) I would be able to read the publications with only minor difficulties.
(04) ( ) I would be able to read the publications with complete confidence and ease.

F. You are planning to take a university course in history of Hispanic peoples of the Southwest. The textbooks for the course are all written in Spanish.

MY READING SKILLS IN SPANISH ARE SUCH THAT: (Please check one)

(01) ( ) I would be able to read the textbooks.
(02) ( ) I would be able to read the textbooks with limited facility.
(03) ( ) I would be able to read the textbooks with only minor difficulties.
(04) ( ) I would be able to read the textbooks with complete confidence and ease.

G. A close friend of yours, living in another part of the state, speaks primarily Spanish. You need to write a note in Spanish inviting her/him to spend a weekend with you.

MY WRITING SKILLS IN SPANISH ARE SUCH THAT: (Please check one)

(01) ( ) I would be unable to write such a note in Spanish.
(02) ( ) I would be able to write such a note in Spanish with only limited facility.
(03) ( ) I would be able to write such a note in Spanish with only minor difficulties.
(04) ( ) I would be able to write such a note in Spanish with complete confidence and ease.

H. You have developed a cultural unit on the Mexican cultural custom of Las Posadas. You need to write a description of this celebration in Spanish to hand out to the parents of your students (one or two pages).

MY WRITING SKILLS IN SPANISH ARE SUCH THAT: (Please check one)

(01) ( ) I would be unable to write such a description in Spanish.
(02) ( ) I would be able to write such a description in Spanish with only limited facility.
(03) ( ) I would be able to write such a description in Spanish with only minor difficulties.
(04) ( ) I would be able to write such a description in Spanish with complete confidence and ease.

I. You are taking a university course on Spanish reading methodology. You have to write a term paper (10-15 pages) in Spanish on some aspect of teaching Spanish reading to bilingual students.

MY WRITING SKILLS IN SPANISH ARE SUCH THAT: (Please check one)

(01) ( ) I would be unable to write such a term paper in Spanish.
(02) ( ) I would be able to write such a term paper in Spanish with only limited facility.
(03) ( ) I would be able to write such a term paper in Spanish with only minor difficulties.
(04) ( ) I would be able to write such a term paper in Spanish with complete confidence and ease.

15. THE CHART BELOW CONTAINS A LIST OF AREAS OF STUDY COVERED IN COLLEGE/UNIVERSITY AND INSERVICE COURSES OFFERED TO PREPARE TEACHERS TO TEACH STUDENTS OF LIMITED ENGLISH-SPEAKING ABILITY. FOR EACH AREA OF STUDY LISTED BELOW, ANSWER QUESTIONS A AND B. INCLUDE ONLY THOSE COURSES COMPLETED BY THE END OF THE 1981-82 ACADEMIC YEAR.

AREA OF STUDY	A. Was the listed area of study covered in your college/university training (courses taken for credit)		B. Was the listed area of study covered in your Inservice training (non-college credit courses)	
	(1)	(2)	(1)	(2)
(01) Teaching the Spanish language arts (including reading) to students whose native language is Spanish	( ) Yes	( ) No	( ) Yes	( ) No

	College/University (1)	(2)	Inservice (1)	(2)
(02) Teaching other subject areas (math, science, etc.) to students whose native language is Spanish	( ) Yes	( ) No	( ) Yes	( ) No
(03) History and culture or ethnic studies associated with the background of students whose native language is Spanish	( ) Yes	( ) No	( ) Yes	( ) No
(04) Teaching English as a second language	( ) Yes	( ) No	( ) Yes	( ) No
(05) Bilingualism and the theory of bilingual education	( ) Yes	( ) No	( ) Yes	( ) No
(06) The study of Spanish solely for the purpose of acquainting teachers with the problems of using a language other than one's mother tongue	( ) Yes	( ) No	( ) Yes	( ) No
(07) Tests and measurement, focused on students whose language is other than English	( ) Yes	( ) No	( ) Yes	( ) No
(08) Guidance and counseling, focused on students whose language is other than English	( ) Yes	( ) No	( ) Yes	( ) No
(09) Materials and curriculum development for bilingual education programs	( ) Yes	( ) No	( ) Yes	( ) No
(10) Linguistics	( ) Yes	( ) No	( ) Yes	( ) No
(11) The theory of learning, specifically related to students whose language is other than English	( ) Yes	( ) No	( ) Yes	( ) No
(12) Language courses for teachers of students speaking languages other than English, e.g., Spanish for teachers of Spanish-speaking children	( ) Yes	( ) No	( ) Yes	( ) No
(13) Other related areas of study (Specify) ______________________	( ) Yes	( ) No	( ) Yes	( ) No

16. HOW MANY SEPARATE COURSES HAVE YOU TAKEN, COVERING ALL OF THE AREAS OF STUDY THAT YOU CHECKED IN:

Question 15A (College/University)

(01) ( ) 0 - 2 courses
(02) ( ) 3 - 4 courses
(03) ( ) 5 - 6 courses
(04) ( ) 7 - 8 courses
(05) ( ) 9 or more

Question 15B (Inservice)

(01) ( ) 0 - 2 courses
(02) ( ) 3 - 4 courses
(03) ( ) 5 - 6 courses
(04) ( ) 7 - 8 courses
(05) ( ) 9 or more

YOU HAVE COMPLETED THIS QUESTIONNAIRE. THANK YOU FOR YOUR COOPERATION.

PLEASE INSERT YOUR COMPLETED QUESTIONNAIRE IN THE ENCLOSED PREADDRESSED ENVELOPE, SEAL, AND RETURN TO ______________________ OF THE SEDL STAFF.

APPENDIX C

Reading and Mathematics Observation System:

Event Form
Rating Summary Sheet
Master Code Sheet
Definition of Codes
Definition of Scales

READING CURRICULUM MATERIALS

OTHER INSTRUCTORS

TIME | GROUP | INSTRUCTOR | CONTENT | READING | ACTIVITY/TASK | RESPONSE

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35

RAPIDS/SEDL Event Form

720 BEST COPY AVAILABLE

RAMOS/SEDL Rating Summary

Observer ______________________ Date ______________________

School ______________________ Teacher ______________________

I. Classroom Organization

A. Class structure:

_____Self-contained class

_____Team taught class

_____Cross-graded class

_____Other. Explain ______________________

____

B. Classroom organization for teaching reading:

_____As a whole

_____In groups

_____As individuals

_____Other. Explain ______________________

____

C. If grouping is used, basis for grouping:

_____By skills

_____By ability

_____By language type

_____Other. Explain ______________________

____

Classroom Organization (continued)

D. If grouping is used, amount of reading instruction devoted to group work:

_____Students spent most of time alloted to reading in groups of 5 or more

_____Students spent more than half of time allotted to reading in groups of 5 or more

_____Students spent less than half of time allotted to reading in groups of 5 or more

_____Students spent most of the time allotted to reading working by themselves or in groups of 4 or less

E. Location(s) of reading activities (check and order according to amount of use):

_____Student seats

_____Grouping of chairs

_____Chalkboard

_____Table

_____Floor

_____Desk of teacher

_____Other. Explain ______________________________

II. <u>Instructional Style</u>

A. Teacher's classroom managerial style (check one).

_____Informal

_____Highly structured

_____Intermediate

B. Pacing of activities:

_____Fast

_____Relaxed

_____Intermediate

C. Teacher's primary role during reading activities:

_____Instruction

_____Facilitation

_____Both of the above (in equal amounts)

_____Other. Explain ______________________________

____

Instructional Style (continued)

D. Teacher-student interaction pattern during reading instruction:

_____Students initiated most interaction

_____Students initiated more than half of the interaction

_____Teacher initiated more than half of the interaction

_____Teacher initiated most interactions

E. Directed and independent work:

_____Teacher instructed students for more than half the total time allotted to reading

_____Teacher instructed students about half the time, and students work independently the rest of the time allotted to reading

_____Students worked independently more than they were instructed by the teacher in reading

_____Students worked independently during most of the time allotted to reading

F. Instructional technique used by teacher during reading instruction would be best characterized as:

(Check one)

_____Parts-to-whole

_____Whole-to-parts

_____Both of the above (in equal amounts)

(Check one)

_____Inductive

_____Deductive

_____Both of the above (in equal amounts)

(Check one)

_____Structured/planned

_____Unstructured/planned

_____Both of the above (in equal amounts)

III. Feedback Pattern

A. Student reinforcement during reading instruction:

_____Students were frequently praised by the teacher

_____Students were occasionally praised by the teacher

_____At least one or two students were praised by the teacher

_____No student was praised by the teacher

Feedback Pattern (continued)

B. Corrective feedback provided during reading instruction:

_____Student errors were usually followed by corrective feedback

_____Student errors were followed by corrective feedback less than half the time

_____Student errors were seldom followed by corrective feedback

_____There were very few student errors

C. Extrinsic motivation

_____Most students did their work in reading without needing to be reminded, rewarded, or otherwise specifically motivated by the teacher

_____Most students did their work in reading about half the time without needing to be reminded, rewarded, or otherwise specifically motivated by the teacher

_____Almost half the students required considerable specific motivation by the teacher to do their work in reading

_____Most of the students required considerable specific motivation by the teacher to do their work in reading

IV. <u>Response Pattern</u>

A. Attention/involvement (interest level):

_____Students spent most of time allotted to reading attending to or carrying out the assigned task

_____Students spent more than half the time allotted to reading attending to or carrying out the assigned task

_____Students spent less than half the time allotted to reading attending to or carrying out the assigned task

_____Students spent little of the time allotted to reading attending to or carrying out the assigned task

B. Response rate and quality (productivity):

_____Students usually gave a lot of correct responses

_____Students gave relatively few responses and these were usually correct

_____Less than half the students' responses were correct, but they gave a lot of responses

_____Students gave relatively few responses and less than half the responses were correct

C. Regulation and self-control:

_____Students showed reasonable self-control behavior most of the time

Regulation and self-control (continued)

_____Students were generally in control of themselves, and there was little need for adult assistance

_____Students lacked self-control to a noticeable extent, leading to occasional adult intervention

_____Students frequently failed to control themselves requiring considerable adult intervention

1 -
4 TIME
24 - hour clock
0005 General comment
0005 to
0000 Field-specific comment
4000 End file

5 TARGET STUDENT

6 STATUS
N Normal
F Out of focus
M Momentary event
D Different
R Return to group
A Adults associated

[GROUP]

7 GROUP ID NUMBER
8 -
9 NUMBER IN GROUP
01 - 99

[INSTRUCTOR]

10 INSTRUCTOR ID NUMBER

11 CLASSIFICATION
A Aide
C Cross-age tutor
D Administrator
N Intern teacher
P Peer tutor
R Resource teacher
S Substitute teacher
T Teacher
U Other adult not usually in the room
V Volunteer

12 ROLE
A Assess/diagnose
C Control
D Discuss
F Facilitate
L Lecture
M Manage class
N Instruction
P Prepare
R Read
S Show
V Observe

RAMOS/SEDL Master Code Sheet

[CONTENT]

13 SUBJECT
A Art
B Class business
C Science
F Free choice
H Spanish
L Language arts
M Mathematics
P Playtime, recess
R Reading
S Social studies
U Music
W Wait time

14 -
16 INSTRUCTIONAL FOCUS
S General skills
SAL Alphabetizing
SAD Auditory discrimination
SCA Composing aloud
SCL Cap-lower case corres.
SDU Dictionary usage
SDW Dictating word
SLN Letter-name work
SOR Oral reading fluency
SPD Phonetic discrimination
SSA Skimming ability
SSL Syllabification
SSP Spelling practice
SSR Silent reading fluency
SVD Visual discrimination
SWL Writing letters
SWP Writing phrases
SWS Writing sentences
SWW Writing words

G Grammar skills
GAD Adjective, adverb, etc.
GCN Contractions
GCP Capitalization
GLW Compound words
GND Noun derivative
GPU Punctuation usage
GSP Singular/plural
GSS Sentence structure

GTU Tense usage
GVD Verb derivation

V Vocabulary skills
VAS Antonyms/synonyms
VEC Easily confusable words
VVE Vocabulary enrichment

14 -
16 (Cont'd)

D Decoding strategies
DCR Clause recognition
DCS Letter cluster - sound recognition
DLR Letter recognition
DLS Letter-sound recognition
DME Morpheme element recognition
DPR Phrase recognition
DSP Spelling pattern recognition
DSR Simple sentence recognition
DWT Whole text recognition
DWW Whole word recognition

C Comprehension strategies
CLF Literal facts
CMD Major ideas
CMI Making inferences
CRE Relations
CSS Story sequence
CTA Talking about what read

N Interpretation strategies
NAH Appreciating humor
NEA Emotional attitude
NPE Predicting events
NSM Sensory images
NSR Seeing relationships

17 TECHNIQUE
P Parts-to-whole
W Whole-to-parts

18 LANGUAGE OF INSTRUCTION
S Spanish
E English
A Alternate use of both
C Code switching

[MATERIALS]

19 -
22 LC Letter card
PC Phonics card
PH Phrase card
PR Picture card

19 -
22 (Cont'd)
SC Sentence card
TC Task card
WC Word card

AW Another workbook
BW Basal workbook
DS Ditto sheet
PT Programmed text
PW Phonics workbook

BT Basal test
CT Commercial test
ST Standardized test
TT Teacher-made test
RF Informal Reading Inventory

BR Basal reader
LB Library Book
SB Supplementary book

AM Art material
AS Auditory stimulus
AV Audio-visual aid
CB Chalkboard
CC Commercial chart
CH Chart - teacher-made
GA Game
LM Language master
MN Magazine, newspaper
PP Paper, pencil
QT Questions from teacher
SM Student-made materials
SO (Other) stuff
TL Teacher manual
TM Teacher-made materials
TW Typewriter
TR Tape recorder

[ACTIVITY TASK]

23 -
25 A Act'y reading related
AA Art activity
DA Dramatization activity
EW Evaluating work
MA Music activity
MC Making correspondences
PA Picture activity
PG Playing games
RI Reading instructions
RV Reciting/responding verse

L Lecture
LL Listen lecture
LS Listen story

D Discussion
SP Speak

23 -
25 (Cont'd)

W Independent work
CM Copying material
CR Creating writing
LD Listen, do
RS Reading silently
TA Teaching another
WA Writing answers
WP Writing practice

Q Question, answers
SA Spoken answers
TT Test taking
WA Written answers
WD Writing from dictation

R Recitation, reading
CA Composing aloud
DR Drill
DS Dictating story
RA Reading aloud
RR Reading and responding
RU Reading in unison

V Audio-visual
WL Watch, listen

T Transitional activity
CU Clean-up
CW Completing work
PW Preparing work
WT Wait time

N Not reading

[RESPONSE]

26-
27 Number non-engaged students 00-99

28-
29 Column rating codes
H High
M Medium
L Low
N None
X Not available

SPECIAL CODES
X Stop (not available)
? Uncertain

8.81

BRS

BEST COPY AVAILABLE

Definition of Codes

INSTRUCTIONAL FOCUS

INSTRUCTIONAL FOCUS CODES

S GENERAL SKILLS

These are general READING SKILLS, not classifiable as grammar, vocabulary, decoding, comprehension, interpretation.

SAL Alphabetizing

SAD Auditory Discrimination - Kindergarten "skills which are nonphonetic."

SCA Composing Aloud - Dictating captions, part of a group story, or an individual story or sentence.

SCL Capital-Lower Case Correspondence

SDU Dictionary Usage - Use of pronunciation key in a dictionary or practice in locating words. For beginning readers, use of a picture dictionary to identify unknown words.

SDW Dictating Words - To the teacher or to a classmate.

SLN Letter-Name Work - Naming letters of the alphabet.

SOR Oral Reading Fluency - Reading for expression.

SPD Phonetic Discrimination - Distinguishing between phonetic contrasts, hearing likenesses and differences among letter sounds as they occur in words. Detecting words that begin or end with the same sound that rhyme, or that contain a given sound where the emphasis is on recognition of the sound are three examples.

SSA Skimming Ability - Scanning reading materials to discover main ideas or to locate details (i.e., locating specific information).

SSL Syllabification - Dividing words into syllables.

SSP Spelling Practice - Emphasis on correct spelling, rather than simply writing words.

SSR Silent Reading Fluency

727

General Skills (continued)

SVD Visual Discrimination - Perception tasks such as matching shapes that are the same, determining similarities of word parts, and task associated with reading readiness.

SWL Writing Letters - For practice or review, not to introduce new concepts.

SWP Writing Phrases - For practice or review, not to introduce new concepts.

SWS Writing Sentences - For practice or review, not to introduce new concepts.

SWW Writing Words - For practice or review, not to introduce new concepts.

G GRAMMAR SKILLS

These are skills that deal with the relationship of words in sentences, word order, and variation in form to show tense, number, change in usage.

GAD Adjective/Adverb, etc. - Form of noun and verb determiners (adjectives and adverbs) and function words (articles, prepositions, conjunctions). Includes concordance with number and gender when appropriate.

GCN Contractions - Shortened form of a single word, or of two words, as can't for cannot or I'll for I will.

GCP Capitalization - The rules determining when words begin with capital letters, e.g., proper nouns or beginning of a sentence, etc.

GCW Compound Words - Words constructed by combining two or more different words, such as housetop, toothbrush

GND Nouns (Derivative) - Sets of word denoting persons, places, things, and noun derivatives such as hardness, noun form of hard; diminutives.

GPU Punctuation Usage - Use of period, comma, question mark, etc., to clarify meaning.

GSP Singular-Plural - Changes in form from singular to plural (horse, horses) and agreement of pronoun(s) with referent(s).

GSS Sentence Structure - Learning the rules for construction of a well-formed sentence, e.g., necessary components such as: subject + verb + object.

GTU Tense Usage - Understanding of proper usage of verbs to represent past, present, future tense, as well as mood (subjunctive/indicative) and aspect.

GVD Verb Derivation - Changes in verbs to show tense, both those derived regularly (ed ending for past tense, walk, walked) and irregularly (is, was/are, were); person-number (walk, walks), mood and aspect.

V VOCABULARY SKILLS

These are skills which sharpen and expand a student's understanding of the meaning and use of words.

VAS Antonyms/Synonyms - Words that are opposed in meaning (antonyms), as good/bad, up/down, or nearly the same in meaning (synonyms), as joyful/glad.

VEC Easily Confusable Words - Words that may be confused by similarity in pronunciation (homophones), as threw, through, or of the same written form but different meaning (homographs), as lead (metal) and lead (to precede).

VVE Vocabulary Enrichment - Experience with the meaning of words, learning new words, and varied meanings for words.

D DECODING SKILLS

These are skills a student uses in approaching the task of recognition or decoding of a visual array. They are defined here by the size and nature of the language unit on which the student focuses her/his attention during the reading act.

DCR Clause Recognition - Drawing upon context, reads by clauses (independent and dependent) rather than by individual words or sentences.

DCS Letter Cluster - Sound Recognition - Relating sound to common letter clusters (consonant blends, diagraphs, syllables) to sound out or decode the word, as distinguished from reading the word as a whole word.

DLR Letter Recognition - Relating letter name to letter form. For example, in reading the word "no" the child will pronounce the letter names n-o until she/he grasps the pronunciation of the word (or syllable, in the case of multisyllabic words m-a, ma; m-a, ma; m-a-m-a, mamá).

DLS Letter - Sound Recognition - Relating sound of the letter, not the letter name, to the letter form to produce (sound out) the word or word part.

DME Morpheme Element Recognition - Recognizing inflectional elements which produce variation in the word form to show tense, number, change in usage, etc. (walk, walking; está, están; trabajar, trabajador; valid, invalid).

Decoding Skills (continued)

DPR Phrase Recognition - Reading by phrases rather than by individual words, clauses, or sentences (with the ball; flying high; con la pelota; a jugar).

DSP Spe. ..ng Pattern Recognition - Relating sound to phonogram (-ight; -and). Recognizing frequently-occurring spelling patterns as the basis for decoding words. A phonogram is defined as a succession of orthographic letters that occurs with the same phonetic value in several or many words.

DSR Simple Sentence Recognition - Draws upon previous knowledge of the structure of the language, uses context clues and/or previous knowledge of the topic to gain meaning from the visual array. Attention to sentence elements is minimal.

DWT Whole Text Recognition - Samples the visual array for distinctive features and relies heavily on context clues, previous life experiences, and knowledge of the language to gain meaning from the text Attention to individual words or other sentence elements is minimal.

DWW Whole Word Recognition - Recognizing words by their distinctive features. Attention to individual word parts is minimal.

C COMPREHENSION SKILLS

These are skills by which a student extracts meaning from groups of words, a single sentence, a paragraph, or an entire story.

CLF Literal Facts - Remembering specific information in a text, or in other kinds of written material. This code applies when the answer to a question can be located within the written material available to the students, and does not reflect a student's own experience.

CMD Major Idea - Synthesizing and/or tying together the major elements in a passage in order to describe the main theme or give the main idea. (Creating or selecting the appropriate title for a story and identifying the major characters in a story are two skills in this category.)

CMN Making Inferences - Applying reason to detail and events in the story in order to derive additional information and understanding about events or characters.

CRE Relations - Recognizing similarities and differences, cause-effect relations, general-to-specific relations, and relations involving comparisons. All of the information needed to establish the relations should be available within the text materials.

CSS Story Sequence - Recognizing the order in which events have occurred in a story, with some understanding of why one thing leads to another. This category is similar to Relations, but should be applied when a story line is especially important and instruction focuses on the plot structure of a story as opposed to simpler cause-effect relationships embedded in the passage. 730

Comprehension Skills (continued)

CTA <u>Talking About What You've Read</u> - General discussion of a story including questions on many different levels, going from literal facts to evaluation. This category is appropriate when the discussion shifts rapidly from one level to another, as from specific facts to main idea to story sequence.

<u>N</u> <u>INTERPRETATION SKILLS</u>

These are skills by which a student shows an understanding of what he has read by some kind of interpretation in a broader context.

NAH <u>Appreciating Humor</u> - Demonstrating an understanding of the humor in a story by explanation or other kind of interpretation.

NEA <u>Emotional Attitude</u> - Identifying, discussing, or describing feelings generated by a story; relating to them on the basis of personal experience(s).

NPE <u>Predicting Events</u> - Predicting and extending events in a story.

NSM <u>Sensory Images</u> - Demonstrating an understanding of the comparisons and figurative language that appeals to the senses.

NSR <u>Seeing Relationships</u> - Perceiving a similarity between something he has read and previous knowledge, another story, or a prior experience. (Note that CRE, <u>Relations</u>, means the relations are determined within the text materials.)

<u>MATERIAL CODE</u>

LC <u>Letter Card</u>

PC <u>Phonics Card</u>

PH <u>Phrase Card</u>

PR <u>Picture Card</u>

SC <u>Sentence Card</u>

731

Material (continued)

TC Task Card

WC Word Card

AW Another Workbook

BW Basal Workbook

DS Ditto Sheet

PT Programmed Text

PW Phonics Workbook

BT Basal Test

CT Commercial Test

ST Standardized Test

TT Teacher-Made Test

BR Basal Reader

LB Library Book

SB Supplementary Book

AM Art Materials

AS Auditory Stimulus

AV Audio-Visual Aid

CB Chalkboard

CC Commercial Chart

CH Chart - Teacher-Made

GA Game

LM Language Master

MN Magazines, Newspapers

PP Paper, Pencil

QT Questions - from Teacher

SM Student-Made Materials

SN (Other) Stuff Not Classified

732

Material (continued)

TL Teacher Manual

TM Teacher-Made Materials

TW Typewriter

TR Tape Recorder

ACTIVITY/TASK CODES

A__ OTHER ACTIVITIES·READING-RELATED

_AA Art Activity - Using art materials such as crayons, scissors, and paste to complete a worksheet or workbook exercise or other reading activity. This activity could have many forms, for example: drawing pictures to illustrate a story, coloring figures that have the same beginning sound, cutting out alike words from a list and pasting them in a row.

_DA Dramatization Activity - Interpreting a story, an event, or a character.

_EW Evaluating Work - Checking work for completeness or accuracy with or without the instructor's supervision.

_MA Music Activity - Participating in a reading-related music task.

_MC Making Correspondences - Finding relationships, perceiving contrasts, or making comparisons.

_PA Picture Activity - Using pictures to identify items, to find similarities, to interpret events, to arrange events sequentially, or to make inferences.

_PG Playing Games

_RN Reading Instructions - Reading ("how to") instructions provided by a game, a workbook, a mimeographed exercise, or an instructor relevant to a forthcoming activity.

_RV Reciting/Responding to Verse, riddle, nursery rhyme.

L__ LECTURE

_LL Listen Lecture - Listening to an oral presentation by the instructor or anyone designated to perform that task.

_LS Listen to Story - Listening to an oral presentation of a story--fiction, nonfiction, language experience, poetry, or spontaneous creation.

733

D__ DISCUSSION

_SF Speak - Saying something at one time or another--in free-form discussion and/or in regular routine of turns.

W__ INDEPENDENT WORK

_CM Copying Material - Copying written material onto paper or the chalkboard.

_CR Creative Writing - Writing original thoughts.

_LD Listen, Do - Listening to an oral stimulus (tape recorder or instructor) and following instructions independently, such as writing answers on a worksheet.

_RS Read Silently - Reading silently in a reading group or independently.

_TA Teaching Another - helping another student, presenting word cards, listening to oral reading, or assisting in the completion of a worksheet.

_WA Written Answers

_WP Writing Practice - Practicing how to form letters.

Q__ QUESTIONS, ANSWERS: Students are given direct questions and are expected to give direct answers. This activity differs from Independent Work in that it is teacher-led. This code is also applicable where students are free to raise questions with the instructor. If the observer is uncertain, the D code should be used to describe a Discussion with some questions and answers.

_SA Spoken Answers

_TT Test-Taking - Writing or speaking answers in a clearly defined test situation.

_WA Written Answers

_WD Writing from Dictation - Writing teacher-directed material.

R__ RECITATION, READING: Reading or reciting aloud as a direct assignment in a group.

_CA Composing Aloud - Composing sentences, phrases, or larger units.

_DR Drill - Practicing by repetition.

_DS Dictating a Story - Dictating a story to the teacher or group.

_RA Reading Aloud

_RR Reading and Responding

_RU Reading in Unison

V__ AUDIO-VISUAL: Viewing an audio-visual presentation such as a movie, record, or TV program, slides or transparencies.

_WL Watch, Listen - Watching and/or listening to an audio visual presentation.

T__ TRANSITIONAL ACTIVITIES: This code is for those times before or after activities or when the instructor leaves the group in the midst of an activity to attend to something else. In an open classroom situation, it may refer to times when a target student waits for correction or help from the instructor before proceeding with his task.

_ U Clean Up

_CW Complete Work - Finishing work and putting away materials.

_PW Preparing to Work - Getting self and/or materials ready for a task.

_WT Waiting Time

N__ NOT READING: Indicates a teacher-directed nonreading activity (as math, science) in which an activity that can be coded is occurring.

LANGUAGE OF INSTRUCTION CODES

S Spanish - Instruction during this event is conducted exclusively in Spanish or primarily in Spanish with only an occasional use of English during an event period.

E English - Instruction during this event is conducted exclusively in English or primarily in English with only an occasional use of Spanish during an event period.

A Alternating use of the two languages - Both languages are used for substantial amounts of time during an event period. This code is used to indicate concurrent use of the two languages (e.g., the instructor conducts an instructional event by presenting a small portion (one, or a few sentences) in Spanish and then repeating the same portion in English. The entire event should be conducted in this matter--shifting back and forth between the two languages with at least one third of the instruction in one or the other language.

This code is also used for alternate use of the two languages (e.g., the instructor presents the entire instructional event in Spanish and then immediately presents the same information in English to the same child or group of children).

As distinguished from code-switching, alternating use of the two languages is characterized by exclusive use of one language at a time during an instructional event.

C <u>Code-Switching</u> - This code is used to describe situations in which the speakers interject into the dominant use of one language long stretches of speech that exhibit the other language's phonological and morphological features. Here are examples of code-switching:

1. No, yo sí brincaba en el trampoline when I was a senior.
(No, I did jump on the trampoline when I was a senior.)

2. La consulta era eight dollars.
(The office visit was eight dollars.)

3. Tenía un vestido que era como de lace.
(I had a dress that looked as if it was made of lace.)

4. Well, I keep starting. Como por un mes todos los días escribo y ya dejo. Last week empecé otra vez.
(Well, I keep starting some. For about a month I write every day and then I stop. Last week I started again.)

5. Me tomé toda la cafetera, the whole coffe pot.
(I drank the whole coffee pot, the whole coffe pot.)

6. And he was laughing 'cause he saw me coming in. Se estaba riendo de mí.
(And he was laughing 'cause he saw me coming in. He was laughing at me.)

Code-switching should not be confused with the process of borrowing. Words such as <u>puchando</u> (from the English verb <u>push</u>) and <u>troca</u> (from the English word <u>truck</u>), which have been assimilated into the Spanish phonological and morphological system, represent borrowing, rather than code-switching. In code-switching all items are used exactly as they are found in the original language.

<u>CLASSIFICATION CODES</u>

A <u>Aide</u> - A paraprofessional who is <u>regularly</u> assigned to a classroom to assist the teacher.

C <u>Cross-Age Tutor</u> - Another student, usually older, who assists in instructional activities.

D <u>Administrator</u> - Any administrator including principals and district staff.

N <u>Intern Teacher</u> - Student teachers or other individuals undergoing supervised training as teachers.

P <u>Peer Tutor</u> - A student in the same class who assists in instructional activities.

R Resource Teacher - A certificated teacher who performs specialized functions in a classroom or school. The area of spe ialization could include reading, mathematics, psychology, art, music, etc.

S Substitute Teacher - A teacher taking the place of the regular teacher, who is absent.

T Teacher - A certificated teacher who is regularly assigned to the class. More than one teacher may be assigned to the class, for example, in a team-taught classroom.

U Other Adult - Other adult not usually in the room.

V Volunteer - An individual, such as a parent or older student, who is not regularly assigned to a classroom, but comes voluntarily on a regular or irregular schedule.

ROLE CODES

A Assess/Diagnose - formal or informal testing or assessment.

B Control - discipline of class or group, or extended control of individual.

C Discussion - talking between teacher and students, or among students. Can be teacher led, but students should be able to initiate their own comments.

F Facilitate - helps students or groups as needed. Students are working independently or in group activities; teacher provides guidance or direction as a result of monitoring or student request.

L Lecture - extended presentation by teacher of academic material. Do not use if teacher is giving directions, or reading orally.

M Manage - class business; giving students non-academic instructions; supervising staff.

N Instruction - direct instruction not fitting under any other heading; includes many classroom activities that are governed by basal texts and management systems.

P Prepare - getting ready to teach; not actively engaged in instruction.

R Read - extended oral reading by teacher, not as part of other instructional activities.

S Show - demonstrating; showing students how something is done, or how something works; may be accompanied by discussion, but demonstration is most significant feature.

V Observing - looking at students' work, monitoring independent or group activities.

RAMOS: Definition of Scales

The scaling of the raw RAMOS codes (see Master Code Sheet and Master Code Definitions for a specification of these) is provided below, giving for each scale, a description of the instructional dimension scaled, and the raw code-numeric code pairing employed. These scalings were used to generate the quality indices of instruction. Quantity indices were computed by taking the ratio of minutes of applicable codes to minutes of student observation time (i.e., for each student, summing the number of minutes during which a given scale contained an applicable raw code, irrespective of its relative quality, and dividing by the total number of minutes the student was observed during the observation period).

Number of Students. Simply the number of students contained in the instructional group.

Instructor Classification. Scaling of the level of the instructor's formal training, ranging from minimal to substantial:

1 = D,U
2 = V
3 = P
4 = C
5 = A
6 = N
7 = S,R,T

Instructor Role. Scaling of the level of formal instruction provided, ranging from non-instructional roles to direct instruction:

1 = P,C,M
3 = V,A
6 = F
7 = R,S
8 = D,L
9 = N

Subject. Scaling of the amount of reading generally required in the subject being taught, ranging from minimal to substantial:

1 = W,P,B,A,U
2 = F
4 = C,M,S
8 = H,L
9 = R

Instructional Focus. Scaling of the relative explicitness of the instructional emphases and strategies employed by the teacher in three instructional subcategories:

Letter-Sound Unit. Scaling of the explicitness of decoding instruction, ranging from work on isolated units to non-explicit letter-sound pairing to explicit letter-sound pairing:

1 = SAD, SCL, SLN, DLR, SAL, GCP, SVD, SWL
2 = DWW, SSP, GCN, GSP, SPD, SWW, SDW, DSR, DWT, SOR, SSR
3 = DCS, DLS, DME, DSP, SSL

Word Unit - Meaning. Scaling of the explicitness of instruction concerning word meaning, ranging from non-explicit to explicit:

1 = SDU
2 = GND, GVD, VEC, GAD, GCW
3 = VAS, VVE, GTU

Sentence and Text Units - Meaning. Scaling of the explicitness of instruction concerning sentence and text meaning, ranging from non-explicit to explicit:

1 = CLF, SSA, SWS, SWP, DCR, DPR
2 = NAH, CSS, CRE, CTA, NPE, NSR, SCA, GPU
3 = CMD, CMN, NEA, NSM, GSS

Technique. Binary scaling of the technique in which skills of visual or auditory pattern recognition are presented, as either parts-to-whole or whole-to-parts:

1 = P
2 = W

Language of Instruction. Scaling of the language usage of the instructor, ranging from all Spanish to all English:

1 = S
2 = C,A
3 = E

Materials (Primary and Ancillary). Scaling of the level of text contained in the materials, ranging from non-text to text materials:

1 = AM, AS, TR, PR
2 = GA, LC, PC, AV, LM, SN
3 = TW, TL, OT
4 = PH, SC, TC, CH, CC, SM, WC
5 = CB, PP, TM
6 = BT, CT, ST, TT, NF
7 = AW, BW, DS, PT, PW
8 = BR, LB, SB, MN

73.

Activity/Task. Scaling of the formal language demand required by the particular activity/tasks the instructor chooses as a vehicle to convey the instructional content in three instructional subcategories:

Non-instructional. Binary scaling of the activity/tasks as either non-instructional or instructional:

1 = TCW, TCU, TPW, TWT
2 = all other activity/task codes

Independent. Scaling of the level of formal language demand for activity/tasks classified as independent work, ranging from minimal to substantial:

1 = AAA, WCM, WMC, WPA, WWP, WAA
2 = QWD, WLD, WRS, WTA, WCW, WWA
3 = QTT, QWA, WCR

Listening and Responding in Group. Scaling of the level of formal language demand for activity/tasks classified as listening and responding in groups, ranging from minimal to substantial:

1 = AMA, ADR, APG, ARU, RDR, RRU, RTA, APA, AMC
2 = VLD, VWL, LLS, RRS, AEW, ARV, AWA, LLD, RLD, RRA, WRA
3 = ARN, LLL, DSP, ACA, ADA, ARR, QSA, RCA, RDS, RRR

Number of Nonengaged Students (Collection Years 3-5 only). Simply the number of students contained in the instructional group which are not engaged in the activity/task assigned.

Productivity. Scaling of the rated productivity of the instructional group, ranging from low to high:

1 = N
2 = L
3 = M
4 = H

Noise. Scaling of the rated noise level of the instructional group (relative to the activity/task required), ranging from low to high:

1 = N
2 = L
3 = M
4 = H

Attention (Collection Years 1-2 only). Scaling of the rated attention of the instructional group relative to the activity/task required, ranging from low to high:

1 = N
2 = L
3 = M
4 = H

APPENDIX D

Reading and Mathematics Observation System:

Sample Coding Analyses for Individual Students and Groups

Table 1

Variable Positions in PAMOS Expanded Raw Coding Analyses

COLUMNS	VARIABLE
1-4	STUDENT ID (Student analyses only)
5-7	TEACHER ID
9	GROUP ID (group analyses only)
10-15	DATE OF OBSERVATION
[illegible]	MINUTES
23	STATUS
24	GROUP ID
25-26	NUMBER OF STUDENTS
27	INSTRUCTOR: ID
28	INSTRUCTOR: CLASSIFICATION
29	INSTRUCTOR: ROLE
30	SUBJECT
31-33	INSTRUCTIONAL FOCUS
34	TECHNIQUE
35	LANGUAGE OF INSTRUCTION
36-37	PRIMARY MATERIALS
38-39	ANCILLARY MATERIALS
40-42	ACTIVITY-TASK
43-44	NUMBER NON-ENGAGED
45	PRODUCTIVITY
46	NOISE
47	ATTENTION
50	MOMENTARY CONTROL
[illegible]-54	OBSERVATION LENGTH IN MINUTES
57-59	OBSERVED LENGTH IN MINUTES

Note: This is a duplicate of Table 1 appearing in the text.

742

BEST COPY AVAILABLE

```
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]  [illegible]  8  NJ051TNRSSLFEOT [illegible]
[illegible]  [illegible]  9  NJ051TNRSSLFEOT [illegible]
[illegible]  [illegible]  10 NJ051TRR [illegible]
[illegible]  [illegible]  11 NJ051TRR [illegible] PE [illegible] LL [illegible]
[illegible]  [illegible]  12 NJ051TNR [illegible] PE [illegible]
[illegible]  [illegible]  13 NJ051TFRSOR×EBR [illegible]
[illegible]  [illegible]  14 NJ051TFRSOR×EBR [illegible]
[illegible]  [illegible]  15 NJ051TFRSOR×EBR [illegible]
[illegible]  [illegible]  16 NJ051TNRCLF×EDPOTCSA [illegible]
[illegible]  [illegible]  17 NJ051TFRSOR×EBR×[illegible]
[illegible]  [illegible]  18 NJ051TFRSOR×EBR×[illegible]
[illegible]  [illegible]  19 NJ051TFRSOR×EBR×[illegible]
[illegible]  [illegible]  20 NJ051TFRSOR×EBR×[illegible]
[illegible]  [illegible]  21 NJ051TFRSOR×EBR×[illegible]
[illegible]  [illegible]  22 NJ051TFRSOR×EBR [illegible]
[illegible]  [illegible]  23 NJ051TFRSOR×EBR×[illegible]
[illegible]  [illegible]  24 NJ051TNRCLF×EBROTCSH [illegible]
[illegible]  [illegible]  25 NJ051TNRCLF×EBROTCSP [illegible]
[illegible]  [illegible]  26 NJ051TNRCLF×EBR [illegible]
[illegible]  [illegible]  27 NJ051TFRSOR×EBR [illegible]
[illegible]  [illegible]  28 NJ051TFRSOR×EBR [illegible]
[illegible]  [illegible]  29 NJ050××LSWS×ECBFF [illegible]
[illegible]  [illegible]  30 NJ050××LSWS×ECBF [illegible]
[illegible]  [illegible]  31 NJ050××LSWS×ECBF [illegible]
[illegible]  [illegible]  32 NJ050××LSWS×ECBF [illegible]
[illegible]  [illegible]  33 NJ050××LSWS×ECBF [illegible]
[illegible]  [illegible]  34 NJ050××LSWS×ECBF [illegible]
[illegible]  [illegible]  35 NJ050××LSWS×ECBFFW [illegible]
[illegible]  [illegible]  36 NJ050××LSWS×ECBFF [illegible]
[illegible]  [illegible]  37 NJ050××LSWS×ECBFF [illegible]
[illegible]  012483  38 NJ050××LSWS×ECBFFW [illegible]
[illegible]  [illegible]  39 NJ050××LSWS×ECBFFWCMOOML×
[illegible]  012483  40 NJ050××LSWS×ECBFFWCM [illegible]
[illegible]  [illegible]  41 NJ050××LSWS×ECBFFWCMOOML×
[illegible]  [illegible]  42 NJ050××LSWS×ECBFFWCM [illegible]
[illegible]  [illegible]  43 NJ050××LSWS×ECBFFWCMOOML×
[illegible]  [illegible]  44 NJ050××LSWS×ECBFFWCM [illegible]
[illegible]  [illegible]  45 NJ050××LSWS×ECBFFWCM [illegible]
[illegible]  [illegible]  46 NJ050××LSWS×ECBFFWCM [illegible]
[illegible]  [illegible]  47 NJ050××LSWS×ECBFFWCM [illegible]
[illegible]  012483  48 NJ050××LSWS×ECBFFWCM [illegible]
[illegible]  [illegible]  49 NJ050××LSWS×ECBFFWCM [illegible]
[illegible]  [illegible]  50 NJ050××LSWS×ECBFFWCMOOML
[illegible]  [illegible]  51 NJ050××LSWS×ECBFFWC [illegible]
```

Figure 1. RAMOS expanded raw coding for target student 5109 based on sample protocol.

```
[illegible]
012483   1  [illegible]
         2  [illegible]
012483   3  [illegible]
012483   4  [illegible]
012483   5  [illegible] LSWS [illegible]
012483   6  [illegible] LSWS [illegible]
012483   7  [illegible] LSWS [illegible]
012483   8  NI051TNRSSLPEDT[illegible]OSA[illegible]ML
012483   9  NI051TNRSSLPEDT[illegible]OSA[illegible]
012483  10  NI051TRR[illegible]PE[illegible]LL5[illegible]
012483  11  NI051TRR[illegible]LL5[illegible]
012483  12  NI051TNR[illegible]PE[illegible]OSAOOML
012483  13  NI051TFRSOR[illegible]EBR[illegible]RRAOOML
012483  14  NI051TFRSOR[illegible]EBR[illegible]RRAOOML
012483  15  NI051TFRSOR[illegible]RRA[illegible]ML
012483  16  NI051TNRCLF[illegible]EBROTOSA[illegible]ML
012483  17  NI051TFRSOR[illegible]EBR[illegible]RRA[illegible]ML
012483  18  NI051TFRSOR[illegible]EBR[illegible]RRA[illegible]ML
012483  19  NI051TFRSOR[illegible]EBR[illegible]RRA[illegible]ML
012483  20  NI051TFRSOR[illegible]EBR[illegible]RRAOOML
012483  21  NI051TFRSOR[illegible]EBR[illegible]RRA[illegible]ML
012483  22  NI051TFRSOR[illegible]EBR[illegible]RRAOOML
012483  23  NI051TFRSOR[illegible]EBR[illegible]RRAOOML
012483  24  NI051TNRCLF[illegible]EBROTOSAOOLL
012483  25  NI051TNRCLF[illegible]EBROTOSA[illegible]LL
012483  26  NI051TNRCLF[illegible]EBROTOSAOOLL
012483  27  NI051TFRSOR[illegible]EBR[illegible]RRA[illegible]ML
012483  28  NI051TFRSOR[illegible]EBR[illegible]RRAOOML
012483  29  NI050[illegible]LSWS[illegible]ECBFFWCMOOML
012483  30  NI050[illegible]LSWS[illegible]ECBFFWCMOOML
012483  31  NI050[illegible]LSWS[illegible]ECBFFWCM[illegible]
012483  32  NI050[illegible]LSWS[illegible]ECBFFWCM[illegible]ML
012483  33  NI050[illegible]LSWS[illegible]ECBFFWCM[illegible]ML
012483  34  NI050[illegible]LSWS[illegible]ECBFFWCMOOML
012483  35  NI050[illegible]LSWS[illegible]ECBFFWC[illegible]ML
012483  36  NI050[illegible]LSWS[illegible]ECBFFWCMOOML
012483  37  NI050[illegible]LSWS[illegible]ECBFFWCMOOML
012483  38  NI050[illegible]LSWS[illegible]ECBFFWCMOOML
012483  39  NI050[illegible]LSWS[illegible]ECBFFWCMOOML
012483  40  NI050[illegible]LSWS[illegible]ECBFFWCMOOML
012483  41  NI050[illegible]LSWS[illegible]ECBFFWCMOOML
012483  42  NI050[illegible]LSWS[illegible]ECBFFWCMOOML
012483  43  NI050[illegible]LSWS[illegible]ECBFFWCMOOML
012483  44  NI050[illegible]LSWS[illegible]ECBFFWCMOOML
012483  45  NI050[illegible]LSWS[illegible]ECBFFWCM[illegible]
012483  46  NI050[illegible]LSWS[illegible]ECBFFWCM[illegible]ML
012483  47  NI050[illegible]LSWS[illegible]ECBFFWC[illegible]ML
012483  48  NI050[illegible]LSWS[illegible]ECBFFWCM[illegible]ML
012483  49  NI050[illegible]LSWS[illegible]ECBFFWCMO[illegible]ML
012483  50  NI050[illegible]LSWS[illegible]ECBFFWCMOOML
012483  51  NI050[illegible]LSWS[illegible]E[illegible]L
```

Figure 2. RAMOS expanded raw coding for target student 5120 based on sample protocol.

[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	4 [illegible]
[illegible]	012483	5 [illegible]
[illegible]	[illegible]	6 N2061TNLGCWXEWCOTFRR[illegible]
[illegible]	[illegible]	7 N2061TNLGCWXEWCOTFRR[illegible]
[illegible]	012483	8 N2060XXLSWSXECBPFWCM[illegible]
5259202	012483	9 N2060XXLSWSXECBPFWCMOOMLX
[illegible]	[illegible]	10 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	11 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	12 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	13 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	14 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	15 N2060XXLSWSXECBPFWCMOOMLX
[illegible]	012483	16 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	17 N2060XXLSWSXECBPFWCMOOMLX
[illegible]	012483	18 N2060XXLSWSXECBPFWCM[illegible]
5259202	012483	19 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	[illegible]	20 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	21 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	22 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	23 N2060XXLSWSXECBPFWCMOOMLX
5259202	012483	24 N2060XXLSWSXECBPFWCMOOMLX
5259202	012483	25 N2060XXLSWSXECBPFWCMOOMLX
[illegible]	012483	26 N2060XXLSWSXECBPFWCMOOML[illegible]
[illegible]	012483	27 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	28 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	29 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	30 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	31 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	32 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	33 N2060XXLSWSXECBPFWCMOO[illegible]
[illegible]	012483	34 N2060XXLSWSXECBPFWCMOOMLX
[illegible]	012483	35 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	36 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	37 N2060XXLSWSXECBPFWCMOOML[illegible]
[illegible]	012483	38 N2060XXLSWSXECBPFWC[illegible]
[illegible]	012483	39 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	40 N2060XXLSWSXECBPFWCM[illegible]
[illegible]	012483	41 N2060XXLSWSXECBPFWCMOOMLX
[illegible]	012483	42 N2060XXLSWSXECBPFWCM[illegible]
5259202	012483	43 N2060XXLSWSXECBPFWCMOOML
[illegible]	012483	44 N2040XXLSWSXECBPFWCM[illegible]
[illegible]	012483	45 N2040XXLSWSXECBPFWCM[illegible]
[illegible]	012483	46 N2040XXLSWSXECBPFWCM[illegible]
[illegible]	012483	47 N2040XXLSWSXECBPFWCMOOML[illegible]
[illegible]	012483	48 N2040XXLSWSXECBPFWCM[illegible]
[illegible]	012483	49 N2040XXLSWSXECBPFWCM[illegible]
[illegible]	012483	50 N2040XXLSWSXECBPFWCM[illegible]
5259202	012483	51 N2040XXLSWSXECBPFWCMOOML[illegible]

Figure 3. RAMOS expanded raw coding for target student 5259 based on sample protocol.

2022	012483	1	N2061TNLDWWXEWCCTRRROOML
2022	012483	2	N2061TNLDWWXEWCCTRRROOML
2022	012483	3	N2061TNLDWWXEWCCTRRROOML
2022	012483	4	N2061TNLDWWXEWCCTRRROOML
2022	012483	5	N2061TNLDWWXEWCCTRRROOMLX
2022	012483	6	N2061TNLSCWXEWCCTRRROOMLX
2022	012483	7	N2061TNLGCWXEWCCTRRROOML
2022	012483	8	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	9	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	10	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	11	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	12	N2060XXLSWSXECBPPWCMOOML
2022	012483	13	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	14	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	15	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	16	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	17	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	18	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	19	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	20	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	21	N2060XXLSWSXECBPPWCMOOML
2022	012483	22	N2060XXLSWSXECBPPWCMOOML
2022	012483	23	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	24	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	25	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	26	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	27	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	28	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	29	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	30	N2060XXLSWSXECBPPWCMOOML
2022	012483	31	N2060XXLSWSXECBPPWCMOOML
2022	012483	32	N2060XXLSWSXECBPPWCMOOML
2022	012483	33	N2060XXLSWSXECBPPWCMOOML
2022	012483	34	N2060XXLSWSXECBPPWCMOOML
2022	012483	35	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	36	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	37	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	38	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	39	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	40	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	41	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	42	N2060XXLSWSXECBPPWCMOOML
2022	012483	43	N2060XXLSWSXECBPPWCMOOMLX
2022	012483	44	N2040XXLSWSXECBPPWCMOOMLX
2022	012483	45	N2040XXLSWSXECBPPWCMOOMLX
2022	012483	46	N2040XXLSWS ECBPPWCMOOML
2022	012483	47	N2040XXLSWSXECBPPWCMOOMLX
2022	012483	48	N2040XXLSWSXECBPPWCMOOMLX
2022	012483	49	N2040XXLSWSXECBPPWCMOOML
2022	012483	50	N2040XXLSWSXECBPPWCMOOML
2022	012483	51	N2040XXLSWSXECBPPWCMOOMLX

Figure 4. RAMOS raw coding for group 2 based on sample protocol.

BEST COPY AVAILABLE

```
2023  012483   1
2023  012483   2
2023  012483   3
2023  012483   4
2023  012483   5
2023  012483   6
2023  012483   7
2023  012483   8 N3051 NRSSLPEDT  QSAOOML
2023  012483   9 N3051TNRSSLPEOTXXQSAOOMLX
2023  012483  10 N3051TRRXXXPEXXXXLLSOOXLX
2023  012483  11 N3051TRRXXXPEXXYXLLSOOXLX
2023  012483  12 N3051TNRXYXPEYXXXQSAOOMLX
2023  012483  13 N3051TFRSORXEBRXXRRAOOMLX
2023  012483  14 N3051TFRSORXEBRXXRRAOOMLX
2023  012483  15 N3051TFRSORXEBRXXRRAOOMLX
2023  012483  16 N3051TNRCLFXEBRQTQSAOOMLX
2023  012483  17 N3051TFRSORXEBRXXRRAOOMLX
2023  012483  18 N3051TFRSORXEBRX RRAOOML
2023  012483  19 N3051TFRSORXEBRXXRRAOOMLX
2023  012483  20 N3051TFRSORXEBRXXRRAOOML
2023  012483  21 N3051TFRSORXEBRXXRRAOOMLX
2023  012483  22 N3051TFRSORXEBRXXRRAOOMLX
2023  012483  23 N3051TFRSORXEBRXXRRAOOMLX
2023  012483  24 N3051TNRCLFXEBRQTQSAOOLLX
2023  012483  25 N3051TNRCLFXEBRQTQSAOOLLX
2023  012483  26 N3051TNRCLFXEBRQTQSAOOLLX
2023  012483  27 N3051TFRSORXEBRXXRRAOOMLX
2023  012483  28 N3051TFRSORXEBRXXRRAOOMLX
2023  012483  29 N3050XXLSWSXECBFFWCMOOMLX
2023  012483  30 N3050XXLSWSXECBFFWCMOOML
2023  012483  31 N3050XXLSWSXECBFFWCMOOMLX
2023  012483  32 N3050XXLSWSXECBFFWCMOOMLX
2023  012483  33 N3050XXLSWSXECBFFWCMOOMLX
2023  012483  34 N3050XXLSWSXECBFFWCMOOMLX
2023  012483  35 N3050XXLSWSXECBFFWCMOOMLX
2023  012483  36 N3050XXLSWSXECBFFWCMOOML
2023  012483  37 N3050XXLSWSXECBFFWCMOOMLX
2023  012483  38 N3050XXLSWSXECBFFWCMOOMLX
2023  012483  39 N3050XXLSWSXECBFFWCMOOMLX
2023  012483  40 N3050XXLSWSXECBFFWCMOOMLX
2023  012483  41 N3050XXLSWSXECBFFWCMOOMLX
2023  012483  42 N3050XXLSWSXECBFFWCMOOMLX
2023  012483  43 N3050XXLSWSXECBFFWCMOOMLX
2023  012483  44 N3050XXLSWSXECBFFWCMOOMLX
2023  012483  45 N3050XXLSWSXECBFFWCMOOMLX
2023  012483  46 N3050XXLSWSXECBFFWCMOOMLX
2023  012483  47 N3050XXLSWSX  BFFWCMOOMLX
2023  012483  48 N3050XXLSWSX  BFFWCMOOMLX
2023  012483  49 N3050XXLSWSXEC3FFWCMOOMLX
2023  012483  50 N3050XXLSWSXECBFFWCMOOMLX
2023  012483  51 N3050XXLSWSXECBFFWCMOOMLX
```

Figure 5. RAMOS raw coding for group 3 based on sample protocol.

BEST COPY AVAILABLE

[illegible]	[illegible]	[illegible]	
[illegible]	[illegible]	[illegible]	
[illegible]	[illegible]	[illegible]	
[illegible]	[illegible]	[illegible]	
2024	012483	5	
2024	012483	6	
2024	012483	7	
2024	012483	8	
2024	012483	9	
2024	012483	10	
2024	012483	11	
2024	012483	12	
2024	012483	13	
2024	012483	14	
2024	012483	15	
2024	012483	16	
2024	012483	17	
2024	012483	18	
2024	012483	19	
2024	012483	20	
2024	012483	21	
2024	012483	22	
2024	012483	23	
2024	012483	24	
2024	012483	25	
2024	012483	26	
2024	012483	27	
2024	012483	28	
2024	012483	29	N4041TNRGCNFECHXXRRROOMLX
2024	012483	30	N4041TNRGCNFECHXXRRROOMLX
2024	012483	31	N4041TNRGCNFECHXXRRROOMLX
2024	012483	32	N4041TNRGCNFECHXXRRROOMLX
2024	012483	33	N4041TNRGCNFECHXXRRROOMLX
2024	012483	34	N4041TNRGCNFECHXXRRROOMLX
2024	012483	35	N4041TNRGCNFECHXXRRROOMLX
2024	012483	36	N4041TNRGCNFECHXXRRROOMLX
2024	012483	37	N4041TNRDCSPEQTXXQSAOOMLX
2024	012483	38	N4041TFRSORXEBRXXRRAOOMLX
2024	012483	39	N4041TFRSORXEBRXXRRAOOMLX
2024	012483	40	N4041TFRSORXEBRXXRRAOOMLX
2024	012483	41	N4041TNRCLFXEBRQTQSAOOMLX
2024	012483	42	N4041TNRCLFXEBRQTQSAOOMLX
2024	012483	43	N4041TNRCLFXEBRQTQSAOOML
2024	012483	44	N4041TNRCLFXEBRQTQSAOOMLX
2024	012483	45	N4041TNRCLFXEBRQTQSAOOMLX
2024	012483	46	N4041TNRCLFXEBRQTQSAOOMLX
2024	012483	47	N4041TNRCLFXEBRQTQSAOOML
2024	012483	48	N4040XXLSWSXECBRRWCMOOMLX
2024	012483	49	N4040XXLSWSXECBRRWCMOOMLX
2024	012483	50	N4040XXLSWSXECBRRWCMOOMLX
2024	012483	51	N4040XXLSWSXECBRRWCMOOMLX

Figure 6. RAMOS raw coding for group 4 based on sample protocol.

Table 2

Variable Positions in RAMOS Expanded Scale Coding Analyses

COLUMNS	VARIABLE
01-04	STUDENT ID (Student analyses only)
05-07	TEACHER ID
08	GROUP ID (group analyses only)
11-16	DATE OF OBSERVATION
20-21	MINUTES
25-26	NUMBER OF STUDENTS
28	INSTRUCTOR: CLASSIFICATION
29	INSTRUCTOR: ROLE
30	SUBJECT
32	INSTRUCTIONAL FOCUS: LETTER-SOUND UNIT
33	INSTRUCTIONAL FOCUS: WORD UNIT
34	INSTRUCTIONAL FOCUS: SENTENCE/TEXT UNIT
36	TECHNIQUE
37	LANGUAGE OF INSTRUCTION
38	PRIMARY MATERIALS
39	ANCILLARY MATERIALS
41	ACTIVITY-TASK: NON-INSTRUCTIONAL
42	ACTIVITY-TASK: INDEPENDENT
43	ACTIVITY-TASK: LISTENING/RESPONDING IN GROUP
45-46	NUMBER NON-ENGAGED
47	PRODUCTIVITY
48	NOISE
49	ATTENTION
52	MOMENTARY CONTROL
55-56	OBSERVATION LENGTH IN MINUTES
59-60	OBSERVED LENGTH IN MINUTES

Note: This is a duplicate of Table 2 appearing in the text.

BEST COPY AVAILABLE

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
5109202	012483	1	19	008	001	0355	210	0320	0
5109202	012483	5	19	008	001	0355	210	0320	0
5109202	012483	6	19	008	001	0355	211	0320	0
5109202	012483	7	19	008	001	0355	210	0320	0
5109202	012483	8	5	799	200	1300	203	0420	0
5109202	012483	9	5	799	200	1300	203	0420	0
5109202	012483	10	5	779	200	1300	202	0020	0
5109202	012483	11	5	779	000	1300	202	0020	0
5109202	012483	12	5	799	200	1300	203	0320	0
5109202	012483	13	5	769	200	0380	202	0320	0
5109202	012483	14	5	769	200	0380	202	0320	0
5109202	012483	15	5	769	200	0380	203	0320	0
5109202	012483	16	5	799	001	0383	203	0320	0
5109202	012483	17	5	769	200	0380	202	0320	0
5109202	012483	18	5	769	200	0380	202	0320	0
5109202	012483	19	5	769	200	0380	202	0320	0
5109202	012483	20	5	769	200	0380	202	0320	0
5109202	012483	21	5	764	200	0380	202	0320	0
5109202	012483	22	5	769	200	0380	202	0320	0
5109202	012483	23	5	769	200	0380	202	0320	0
5109202	012483	24	5	799	001	0383	203	0320	0
5109202	012483	25	5	799	001	0383	203	0320	0
5109202	012483	26	5	799	001	0383	203	0320	0
5109202	012483	27	5	769	200	0380	202	0320	0
5109202	012483	28	5	769	200	0380	202	0320	0
5109202	012483	29	5	008	001	0355	210	0320	0
5109202	012483	30	5	008	001	0355	210	0320	0
5109202	012483	31	5	008	001	0355	210	0320	0
5109202	012483	32	5	008	001	0355	210	0320	0
5109202	012483	33	5	008	001	0355	210	0320	0
5109202	012483	34	5	008	001	0355	210	0320	
5109202	012483	35	5	008	001	0355	210	0320	0
5109202	012483	36	5	008	001	0355	210	0320	0
5109202	012483	37	5	008	001	0355	210	0320	0
5109202	012483	38	5	008	001	0355	210	0320	0
5109202	012483	39	5	008	001	0355	210	0320	0
5109202	012483	40	5	008	001	0355	210	0320	0
5109202	012483	41	5	008	001	0355	210	0320	0
5109202	012483	42	5	008	001	0355	210	0320	0
5109202	012483	43	5	008	001	0355	210	0320	0
5109202	012483	44	5	008	001	0355	210	0320	0
5109202	012483	45	5	008	001	0355	210	0320	0
5109202	012483	46	5	008	001	0355	210	0320	0
5109202	012483	47	5	008	001	0355	210	0320	0
5109202	012483	48	5	008	001	0355	210	0320	0
5109202	012483	49	5	008	001	0355	210	0320	0
5109202	012483	50	5	008	001	0355	210	0320	0
5109202	012483	51	5	008	001	0355	210	0320	0

Figure 7. RAMOS expanded scale coding for target student 5109 based on sample protocol.

730

[illegible]	[illegible]			[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	
	[illegible]		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0355	[illegible]	[illegible]	
[illegible]	[illegible]	[illegible]	19	008	001	0355	210		
[illegible]	[illegible]	5	19	008	001	0355	210	[illegible]	[illegible]
[illegible]	012483	6	19	[illegible]	[illegible]	0355	210	[illegible]	[illegible]
5120202	012483	7	19	008	001	0355	210	0220	
[illegible]	012483	8	5	799	200	[illegible]	202	0420	0
5120202	012483	9	5	799	200	1330	203	0420	0
5120202	012483	10	5	779	000	1300	202	0020	0
5120202	012483	11	5	779	000	1300	202	0020	0
5120202	012483	12	5	788	000	1300	203	0220	0
5120202	012483	13	5	769	200	0380	202	0220	0
5120202	012483	14	5	769	200	0380	[illegible]	0220	[illegible]
[illegible]	012483	15	5	769	200	0380	[illegible]	0220	0
5120202	012483	16	5	799	001	0383	[illegible]	0220	0
5120202	012483	17	5	769	200	0380	[illegible]	0220	0
5120202	012483	18	5	759	200	0380	202	0220	[illegible]
5120202	012483	19	5	769	200	0380	202	0220	[illegible]
[illegible]	012483	20	5	769	200	0380	202	0220	[illegible]
[illegible]	012483	21	5	769	200	0380	202	[illegible]	0
[illegible]	012483	22	5	769	200	0380	202	[illegible]	[illegible]
5120202	012483	23	5	769	200	0380	202	0220	[illegible]
5120202	01 83	24	5	799	001	0383	203	0220	[illegible]
5120202	012483	25	5	799	001	0383	203	0220	0
5120202	012483	26	5	799	001	0383	203	0220	[illegible]
5120202	012483	27	5	769	200	0380	[illegible]	[illegible]	0
5120202	012483	28	5	759	200	0380	202	0220	[illegible]
5120202	012483	29	5	008	001	0355	210	0220	[illegible]
[illegible]	012483	30	5	008	001	0355	210	[illegible]	
[illegible]	012483	31	5	008	001	0355	210	[illegible]	[illegible]
[illegible]	012483	32	5	008	001	0355	[illegible]	[illegible]	0
5120202	012483	33	5	008	001	0355	210	0220	0
[illegible]	012483	34	5	008	001	0355	210	0220	[illegible]
5120202	012483	35	5	008	001	0355	210	0220	[illegible]
[illegible]	012483	36	5	008	001	0355	[illegible]	[illegible]	[illegible]
5120202	012483	37	5	008	001	0355	[illegible]	0220	
5120202	012483	38	5	008	001	0355	210	0220	[illegible]
5120202	012483	39	5	008	001	0355	210	0220	[illegible]
5120202	012483	40	5	006	001	0355	210	0220	[illegible]
5120202	012483	41	5	008	001	0355	210	0220	[illegible]
5120202	012483	42	5	008	001	0355	210	[illegible]	[illegible]
5120202	012483	43	5	008	001	0355	210	0220	0
5120202	012483	44	5	008	001	0355	210	0220	0
5120202	012483	45	5	008	001	0355	210	0220	[illegible]
5120202	012483	46	5	008	001	0355	210	0220	0
5120202	012483	47	5	008	001	0355	210	0220	[illegible]
5120202	012483	48	5	008	001	0355	210	0220	0
5120202	012483	49	5	008	001	0355	210	0220	[illegible]
5120202	012483	50	5	008	001	0355	210	0220	[illegible]
5120202	012483	51	5	008	001	0355	210	0220	[illegible]

Figure 8. RAMOS expanded scale coding for targ [illegible] 5120 based on sample protocol.

BEST COPY AVAILABLE

[illegible]	[illegible]		[illegible]	798	[illegible]	[illegible]	[illegible]	[illegible]	
[illegible]	[illegible]	[illegible]	[illegible]	798	200	034[illegible]	[illegible]	[illegible]	
[illegible]	[illegible]	[illegible]	[illegible]	798	200	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	012483	[illegible]	[illegible]	798	200	0343	207	0320	0
[illegible]	012483	5	6	798	200	0343	207	0320	[illegible]
[illegible]	012483	6	6	798	020	0343	207	0320	[illegible]
[illegible]	012483	7	6	798	020	0343	207	0320	[illegible]
5259202	012483	8	6	008	001	0355	210	0320	0
5259202	012483	9	6	008	001	0355	210	0320	0
5259202	012483	10	6	008	001	0355	210	0320	0
5259202	012483	11	6	008	001	0355	210	0320	0
5259202	012483	12	6	008	001	0355	210	0320	0
5259202	012483	13	6	008	001	0355	210	0320	0
5259202	012483	14	6	008	001	0355	210	0320	0
5259202	012483	15	6	008	001	0355	210	0320	0
5259202	012483	16	6	008	001	0355	210	0320	0
5259202	012483	17	6	008	001	0355	210	0320	0
5259202	012483	18	6	008	001	0355	210	0320	[illegible]
5259202	012483	19	6	008	001	0355	210	0320	[illegible]
5259202	012483	20	6	008	001	0355	210	0320	0
5259202	012483	21	6	008	001	0355	210	0320	0
5259202	012483	22	6	008	001	0355	210	0320	[illegible]
5259202	012483	23	6	008	001	0355	210	0320	0
5259202	012483	24	6	008	001	0355	210	0320	0
5259202	012483	25	6	008	001	0355	210	0320	0
5259202	012483	26	6	008	001	0355	210	0320	[illegible]
5259202	012483	27	6	008	001	0355	210	0320	[illegible]
5259202	012483	28	6	008	001	0355	210	0320	0
5259202	012483	29	6	008	001	0355	210	0320	0
5259202	012483	30	6	008	001	0355	210	0320	0
5259202	012483	31	6	008	001	0355	210	0320	[illegible]
5259202	012483	32	6	008	001	0355	210	0320	[illegible]
5259202	012483	33	6	008	001	0355	210	0320	0
5259202	012483	34	6	008	001	0355	210	0320	
5259202	012483	35	6	008	001	0355	210	0320	[illegible]
5259202	012483	36	6	008	001	0355	210	0320	0
5259202	012483	37	6	008	001	0355	210	0320	0
5259202	012483	38	6	008	001	0355	210	0320	0
5259202	012483	39	6	008	001	0355	210	0320	0
5259202	012483	40	6	008	001	0355	210	0320	0
5259202	012483	41	6	008	001	0355	210	0320	[illegible]
5259202	012483	42	6	008	001	0355	210	0320	[illegible]
5259202	012483	43	6	008	001	0355	210	0320	
5259202	012483	44	4	008	001	0355	210	0320	[illegible]
5259202	012483	45	4	008	001	0355	210	0320	[illegible]
5259202	012483	46	4	008	001	0355	210	0320	[illegible]
5259202	012483	47	4	008	001	0355	210	0320	[illegible]
5259202	012483	48	4	008	001	0355	210	0320	0
5259202	012483	49	4	008	001	0355	210	0320	0
5259202	012483	50	4	008	001	0355	210	0320	
5259202	012483	51	4	008	001	0355	210	0320	[illegible]

Figure 9. RAMOS expandea scale coding for target student 5259 based on sample protocol.

152

BEST COPY AVAILABLE

2022	012483	1	6	798	200	0343	203	0320	0	51	51
2022	012483	2	6	798	200	0343	203	0320	0		
2022	012483	3	6	798	200	0343	203	0320	0		
2022	012483	4	6	798	200	0343	203	0320	0		
2022	012483	5	6	798	200	0343	203	0320	0		
2022	012483	6	6	798	020	0343	203	0320	0		
2022	012483	7	6	798	020	0343	203	0320	0		
2022	012483	8	6	008	001	0355	210	0320	0		
2022	012483	9	6	008	001	0355	210	0320	0		
2022	012483	10	6	008	001	0355	210	0320	0		
2022	012483	11	6	008	001	0355	210	0320	0		
2022	012483	12	6	008	001	0355	210	0320	0		
2022	012483	13	6	00[illegible]	001	0355	210	0320	0		
2022	012483	14	6	008	001	0355	210	0320	0		
2022	012483	15	6	008	001	0355	210	0320	0		
2022	012483	16	6	008	001	0355	210	0320	0		
2022	012483	17	6	008	001	0355	210	0320	0		
2022	012483	18	6	008	001	0355	210	0320	0		
2022	012483	19	6	008	001	0355	210	0320	0		
2022	012483	20	6	008	001	0355	210	0320	0		
2022	012483	21	6	008	001	0355	210	0320	0		
2022	012483	22	6	008	001	0355	210	0320	0		
2022	012483	23	6	008	001	0355	210	0320	0		
2022	012483	24	6	008	001	0355	210	0320	0		
2022	012483	25	6	008	001	0355	210	0320	0		
2022	012483	26	6	008	001	0355	210	0320	0		
2022	012483	27	6	008	001	0355	210	0320	0		
2022	012483	28	6	008	001	0355	210	0320	0		
2022	012483	29	6	008	001	0355	210	0320	0		
2022	012483	30	6	008	001	0355	210	0320	0		
2022	012483	31	6	008	001	0355	210	0320	0		
2022	012483	32	6	008	001	0355	210	0320	0		
2022	012483	33	6	008	001	0355	210	0320	0		
2022	012483	34	6	008	001	0355	210	0320	0		
2022	012483	35	6	008	001	0355	210	0320	0		
2022	012483	36	6	008	001	0355	210	0320	0		
2022	012483	37	6	008	001	0355	210	0320	0		
2022	012483	38	6	008	001	0355	210	0320	0		
2022	012483	39	6	008	001	0355	210	0320	0		
2022	012483	40	6	008	001	0355	210	0320	0		
2022	012483	41	6	008	001	0355	210	0320	0		
2022	012483	42	6	008	001	0355	210	0320	0		
2022	012483	43	6	008	001	0355	210	0320	0		
2022	012483	44	4	008	001	0355	210	0320	0		
2022	012483	45	4	008	001	0355	210	0320	0		
2022	012483	46	4	008	001	0355	210	0320	0		
2022	012483	47	4	008	001	0355	210	0320	0		
2022	012483	48	4	008	001	0355	210	0320	0		
2022	012483	49	4	008	001	0355	210	0320	0		
2022	012483	50	4	008	001	0355	210	0320	0		
2022	012483	51	4	008	001	0355	210	0320	0		

Figure 10. RAMOS expanded scale coding for group 2 based on sample

BEST COPY AVAILABLE

2023	012483	1	-1	000	000	0000	000	-1000	0	51	44
2023	012483	2	-1	000	000	0000	000	-1000	0		
2023	012483	3	-1	000	000	0000	000	-1000	0		
2023	012483	4	-1	000	000	0000	000	-1000	0		
2023	012483	5	-1	000	000	0000	000	-1000	0		
2023	012483	6	-1	000	000	0000	000	-1000	0		
2023	012483	7	-1	000	000	0000	000	-1000	0		
2023	012483	8	5	799	300	1330	203	0420	0		
2023	012483	9	5	799	300	1330	203	0420	0		
2023	012483	10	5	779	000	1300	2[illegible]	0020	0		
2023	012483	11	5	779	000	1300	202	0020	0		
2023	012483	12	5	799	000	1300	203	0320	0		
2023	012483	13	5	769	200	0380	202	0320	0		
2023	012483	14	5	769	200	0380	202	0320	0		
2023	012483	15	5	769	200	0380	202	0320	0		
2023	012483	16	5	799	001	0383	203	0320	0		
2023	012483	17	5	769	200	0380	202	0320	0		
2023	012483	18	5	769	200	0380	202	0320	0		
2023	012483	19	5	769	200	0380	202	0320	0		
2023	012483	20	5	769	200	0380	202	0320	0		
2023	012483	21	5	769	200	0380	202	0320	0		
2023	012483	22	5	769	200	0380	202	0320	0		
2023	012483	23	5	769	200	0380	202	0320	0		
2023	012483	24	5	799	001	0383	203	0220	0		
2023	012483	25	5	799	001	0383	203	0220	0		
2023	012483	26	5	799	001	0383	203	0220	0		
2023	012483	27	5	769	200	0380	202	0320	0		
2023	012483	28	5	769	200	0380	202	0320	0		
2023	012483	29	5	008	001	0355	210	0320	0		
2023	012483	30	5	008	001	0355	210	0320	0		
2023	012483	31	5	008	001	0355	210	0320	0		
2023	012483	32	5	008	001	0355	210	0320	0		
2023	012483	33	5	008	001	0355	210	0320	0		
2023	012483	34	5	008	001	0355	210	0320	0		
2023	012483	35	5	008	001	0355	210	0320	0		
2023	012483	36	5	008	001	0355	210	0320	0		
2023	012483	37	5	008	001	0355	210	0320	1		
2023	012483	38	5	008	001	0355	210	0320	0		
2023	012483	39	5	008	001	0355	210	0320	0		
2023	012483	40	5	008	001	0355	210	0320	0		
2023	012483	41	5	008	001	0355	210	0320	0		
2023	012483	42	5	008	001	0355	210	0320	0		
2023	012483	43	5	008	001	0355	210	0320	0		
2023	012483	44	5	008	001	0355	210	0320	0		
2023	012483	45	5	008	001	0355	210	0320	0		
2023	012483	46	5	008	001	0355	210	0320	0		
2023	012483	47	5	008	001	0355	210	0320	0		
2023	012483	48	5	008	001	0355	210	0320	0		
2023	012483	49	5	008	001	0355	210	0320	0		
2023	012483	50	5	008	001	0355	210	0320	0		
2023	012483	51	5	008	001	0355	210	0320	0		

Figure 11. RAMOS expanded scale coding for group 3 based on sample protocol.

BEST COPY AVAILABLE

2024	012483	1	-1	000	000	0000	000	-1000	0	31	33
2024	012483	2	-1	000	000	0000	000	-1000	0		
2024	012483	3	-1	000	000	0000	000	-1000	0		
2024	012483	4	-1	000	000	0000	000	-1000	0		
2024	012483	5	-1	000	000	0000	000	-1000	0		
2024	012483	6	-1	000	000	0000	000	-1000	0		
2024	012483	7	-1	000	000	0000	000	-1000	0		
2024	012483	8	-1	000	000	0000	000	-1000	0		
2024	012483	9	-1	000	000	0000	000	-1000	0		
2024	012483	10	-1	000	000	0000	000	-1000	0		
2024	012483	11	-1	000	000	0000	000	-1000	0		
2024	012483	12	-1	000	000	0000	000	-1000	0		
2024	012483	13	-1	000	000	0000	000	-1000	0		
2024	012483	14	-1	000	000	0000	000	-1000	0		
2024	012483	15	-1	000	000	0000	000	-1000	0		
2024	012483	16	-1	000	000	0000	000	-1000	0		
2024	012483	17	-1	000	000	0000	000	-1000	0		
2024	012483	18	-1	000	000	0000	000	-1000	0		
2024	012483	19	-1	000	000	0000	000	-1000	0		
2024	012483	20	-1	000	000	0000	000	-1000	0		
2024	012483	21	-1	000	000	0000	000	-1000	0		
2024	012483	22	-1	000	000	0000	000	-1000	0		
2024	012483	23	-1	000	000	0000	000	-1000	0		
2024	012483	24	-1	000	000	0000	000	-1000	0		
2024	012483	25	-1	000	000	0000	000	-1000	0		
2024	012483	26	-1	000	000	0000	000	-1000	0		
2024	012483	27	-1	000	000	0000	000	-1000	0		
2024	012483	28	-1	000	000	0000	000	-1000	0		
2024	012483	29	4	799	200	1340	203	0320	0		
2024	012483	30	4	799	200	1340	203	0320	0		
2024	012483	31	4	799	200	1340	203	0320	0		
2024	012483	32	4	799	200	1340	203	0320	0		
2024	012483	33	4	799	200	1340	203	0320	0		
2024	012483	34	4	799	200	1340	203	0320	0		
2024	012483	35	4	799	200	1340	203	0320	0		
2024	012483	36	4	799	200	1340	203	0320	0		
2024	012483	37	4	799	300	1330	203	0320	0		
2024	012483	38	4	769	200	0380	202	0320	0		
2024	012483	39	4	769	200	0380	202	0320	0		
2024	012483	40	4	769	200	0380	202	0330	0		
2024	012483	41	4	799	001	0383	203	0320	0		
2024	012483	42	4	799	001	0383	203	0320	0		
2024	012483	43	4	799	001	0383	203	0320	0		
2024	012483	44	4	799	001	0383	203	0320	0		
2024	012483	45	4	799	001	0353	203	0320	0		
2024	012483	46	4	799	001	0383	203	0320	0		
2024	012483	47	4	799	001	0383	203	0320	0		
2024	012483	48	4	008	001	0355	210	0320	0		
2024	012483	49	4	008	001	0355	210	0320	0		
2024	012483	50	4	008	001	0355	210	0320	0		
2024	012483	51	4	008	001	0355	210	0320	0		

Figure 12. RAMOS expanued scale coding for group 4 based on sample protocol.

755

APPENDIX E

Reading and Mathematics Observation System - English:

Descriptive Statistics on the Instructional Indices by Instructional Year for Each Site for the Bilingual Sample

Table 1

Reading and Mathematics Observation System - English:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 0 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
NSTUMN	M	10.1	9.6	10.3	13.6	13.8
NSTUMN	S	1.0	2.7	2.5	3.4	4.[illegible]
NSTUMN	N	30	53	50	50	53
CLSSMN	M	5.4	6.3	5.7	6.6	5.7
CLSSMN	S	0.2	0.5	0.4	0.4	0.5
CLSSMN	N	30	53	50	50	53
ROLEMN	M	7.5	7.2	6.8	6.8	5.7
ROLEMN	S	0.6	0.7	1.5	0.9	1.5
ROLEMN	N	30	53	50	50	53
SBJCMN	M	7.5	7.7	8.0	8.1	8.3
SBJCMN	S	0.7	0.8	0.7	0.7	0.7
SBJCMN	N	30	53	50	50	53
IFLTMN	M	1.8	2.2	2.0	2.1	1.9
IFLTMN	S	0.3	0.2	0.3	0.2	0.1
IFLTMN	N	30	53	50	50	53
IFWOMN	M	3.0	2.9	2.7	2.3	2.7
IFWOMN	S	.0	0.2	0.4	[illegible]	0.5
IFWOMN	N	30	57	30		53
IFSTMN	M	2.0	1.6	1.8	1.8	1.6
IFSTMN	S	0.3	0.3	0.4	0.3	0.5
IFSTMN	N	30	53	45	50	53
TECHMN	M	1.0	1.0	1.5	1.0	1.0
TECHMN	S	0.2	0.2	0.3	0.2	0.3
TECHMN	N	30	53	50	50	53
LANGMN	M	2.8	2.9	3.0	3.0	3.0
LANGMN	S	0.3	0.1	.0	.0	0.1
LANGMN	N	30	53	50	50	53
MAT1MN	M	4.1	5.2	5.1	6.8	6.7
MAT1MN	S	1.2	0.9	1.2	0.8	1.0
MAT1MN	N	30	53	50	50	53
MAT2MN	M	2.7	4.2	4.2	4.4	4.5
MAT2MN	S	0.4	1.0	0.9	0.7	0.8
MAT2MN	N	30	53	50	50	53
ATNTMN	M	1.7	1.9	1.9	1.9	1.9
ATNTMN	S	.0	.0	0.1	.0	.0
ATNTMN	N	30	53	30	50	53
ATINMN	M	1.8	1.6	1.8	2.0	2.0
ATINMN	S	0.6	0.3	0.4	0.2	0.3
ATINMN	N	30	53	49	50	52
ATLRMN	M	2.1	2.6	2.3	2.4	2.4
ATLRMN	S	0.4	0.3	0.2	0.3	0.4
ATLRMN	N	30	53	50	50	53

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
MEN2MN	M		3.0	3.4	3.3	3.1
MEN2MN	S		[illegible]	0.8	1.1	1.4
MEN2MN	N		16	30	50	53
PROCMN	M	3.3	3.3	3.3	3.5	3.2
PROCMN	S	0.1	0.3	0.3	0.3	0.6
PROCMN	N	30	53	50	50	53
NOISMN	M	2.8	2.8	2.6	2.4	2.6
NOISMN	S	0.3	0.3	0.3	0.3	0.5
NOISMN	N	30	53	50	50	53
ATTIMN	M	3.7	3.4	3.7		
ATTIMN	S	0.1	0.2	0.2		
ATTIMN	N	30	40	23		
CTRLMN	M	1.0	0.5	0.3	0.2	0.2
CTRLMN	S	0.7	0.5	0.4	0.3	0.4
CTRLMN	N	30	53	50	50	53
OBTMMN	M	51.2	53.4	51.6	57.1	54.1
OBTMMN	S	2.8	5.4	[illegible]	5.4	3.5
OBTMMN	N	30	53	50	50	53
STTMMN	M	45.3	41.9	43.3	48.[illegible]	47.4
STTMMN	S	3.8	8.5	13.2	7.3	5.6
STTMMN	N	30	53	50	50	53

Table 2

BEST COPY AVAILABLE

Reading and Mathematics Observation System - English:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for Site 0 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
CLSSPR	M	80.4	71.0	77.5	68.1	67.3
CLSSPR	S	11.1	19.3	19.8	18.7	21.2
CLSSPR	N	30	53	50	50	33
ROLEPR	M	80.4	70.5	78.0	68.0	66.9
ROLEPR	S	11.1	18.9	20.0	18.6	21.4
ROLEPR	N	30	53	50	50	33
IFLTPR	M	49.0	59.6	55.8	42.2	[illegible].7
IFLTPR	S	9.9	15.7	15.7	17.1	13.5
IFLTPR	N	30	53	50	50	33
IFWDPR	M	18.9	9.4	7.9	15.5	25.2
IFWDPR	S	8.7	10.8	11.2	12.2	19.7
IFWDPR	N	30	53	50	50	33
IFSTPR	M	18.7	16.3	22.2	31.4	38.3
IFSTPR	S	6.3	11.5	14.3	14.1	11.5
IFSTPR	N	30	53	50	50	33
MAT1PR	M	95.0	89.9	88.0	91.8	90.3
MAT1PR	S	2.2	7.3	7.4	5.5	8.4
MAT1PR	N	30	53	50	50	33
MAT2PR	M	73.1	71.5	72.8	75.2	81.0
MAT2PR	S	16.5	10.8	13.6	14.1	12.2
MAT2PR	N	30	53	50	50	33
ATINPR	M	21.0	30.9	28.7	37.2	34.4
ATINPR	S	16.8	16.0	15.0	12.8	18.2
ATINPR	N	30	53	50	50	33
ATLRPR	M	71.2	55.0	56.6	54.2	55.3
ATLRPR	S	15.5	16.1	15.0	13.7	19.2
ATLRPR	N	30	53	50	50	33
IFTTPR	M	86.7	85.3	85.9	89.1	89.2
IFTTPR	S	5.5	8.4	9.4	6.3	4.5
IFTTPR	N	30	53	50	50	33
IFT2PR	M	37.6	25.7	30.1	46.9	63.5
IFT2PR	S	11.4	13.2	14.6	15.0	13.9
IFT2PR	N	30	53	50	50	33
ATTTPR	M	92.2	85.9	85.3	91.4	89.7
ATTTPR	S	2.9	7.2	6.9	4.0	4.4
ATTTPR	N	30	[illegible]	50	50	33
IFLTRP	M	57.1	69.8	65.4	46.8	28.8
IFLTRP	S	12.0	15.3	17.3	18.3	15.2
IFLTRP	N	30	53	50	50	33
IFWDRP	M	54.7	32.2	29.3	31.3	39.1
IFWDRP	S	17.4	28.9	32.6	20.3	15.5
IFWDRP	N	30	53	48	50	33
ATLRRP	M	77.3	63.9	67.0	59.2	61.5
ATLRRP	S	17.9	17.9	16.4	14.0	20.7
ATLRRP	N	30	53	50	50	33
STTMRT	M	94.4	77.0	84.1	97.3	98.3
STTMRT	S	4.7	17.2	17.1	5.5	2.6
STTMRT	N	30	53	50	50	53

Table 7

Reading and Mathematics Observation System - English:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 1 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
NSTDMN	M	12.0	13.7	16.4	18.3	15.4
NSTDMN	S	2.4	3.1	2.6	3.0	4.5
NSTDMN	N	9	12	11	14	9
CLSSMN	M	6.2	6.4	6.7	6.9	6.9
CLSSMN	S	0.5	0.4	0.4	0.3	0.3
CLSSMN	N	6	12	11	14	9
ROLEMN	M	5.7	4.6	4.7	5.0	5.7
ROLEMN	S	1.7	2.4	1.1	1.1	0.9
ROLEMN	N	9	12	11	14	9
OBJCMN	M	7.3	8.0	7.4	7.3	7.0
OBJCMN	S	0.4	0.5	1.4	0.6	0.9
OBJCMN	N	6	12	11	14	9
IFLTMN	M	1.6	1.8	1.8	1.7	2.0
IFLTMN	S	0.3	0.4	0.2	0.1	0.1
IFLTMN	N	6	10	11	14	9
IFWDMN	M	3.0	3.0	2.6	2.6	2.8
IFWDMN	S	0.0	0.1	0.5	0.5	0.1
IFWDMN	N	3	6	8	14	6
IFSTMN	M	2.0	1.2	1.2	1.5	1.8
IFSTMN	S	[illegible]	0.3	0.2	0.4	0.5
IFSTMN	N	5	11	11	14	8
TECHMN	M	1.2	1.4	1.3	1.3	1.3
TECHMN	S	0.1	0.4	0.2	0.2	0.2
TECHMN	N	9	12	11	14	9
LANGMN	M	2.5	3.0	3.0	2.7	2.9
LANGMN	S	0.5	.0	.0	0.2	0.1
LANGMN	N	9	12	11	14	9
MAT1MN	M	4.6	5.1	5.3	6.9	6.5
MAT1MN	S	1.7	0.7	0.3	0.7	0.4
MAT1MN	N	6	12	11	14	9
MAT2MN	M	3.7	4.3	4.5	4.7	4.3
MAT2MN	S	1.5	0.5	0.5	0.6	0.6
MAT2MN	N	6	12	11	14	8
ATNTMN	M	1.9	1.9	1.9	1.9	1.9
ATNTMN	S	.0	0.1	0.2	0.1	0.1
ATNTMN	N	9	12	11	14	9
ATINMN	M	1.6	1.3	1.3	1.9	2.0
ATINMN	S	0.3	0.3	0.7	0.3	0.4
ATINMN	N	5	12	11	14	9
ATLRMN	M	1.4	1.9	2.0	2.2	2.3
ATLRMN	S	0.2	0.5	0.5	0.2	0.3
ATLRMN	N	9	10	11	14	9

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
MENZMN	M		3.1	2.8	2.4	3.3
MENZMN	S		0.1	1.5	1.9	1.7
MENZMN	N		4	11	14	9
PROCMN	M	3.4	2.8	3.1	3.0	3.0
PROCMN	S	0.1	0.5	0.5	0.5	0.1
PROCMN	N	5	12	11	14	9
NOISMN	M	3.0	2.7	2.4	2.3	2.2
NOISMN	S	0.4	0.4	0.5	0.2	0.2
NOISMN	N	9	12	11	14	9
ATTIMN	M	3.1	2.9			
ATTIMN	S	0.3	0.2			
ATTIMN	N	6	3			
CTRLMN	M	0.2	0.3	0.9	0.3	0.2
CTRLMN	S	0.3	0.8	0.6	0.4	0.2
CTRLMN	N	6	12	11	14	9
OBTMMN	M	33.2	29.7	31.7	34.0	36.3
OBTMMN	S	3.1	1.3	[illegible]	4.2	3.2
OBTMMN	N	9	12	11	14	9
STTMMN	M	33.4	29.7	31.7	32.2	31.3
STTMMN	S	5.7	1.8	1.3	3.8	[illegible]
STTMMN	N	9	12	11	14	9

Table 4

Reading and Mathematics Observation System - English: Descriptive Statistics on the Quantity Indices of Instruction by Instructional Year for Site 1 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
CLSSPR	M	100.0	49.1	52.0	61.8	57.6
CLSSPR	S	0.0	28.2	18.8	15.1	18.4
CLSSPR	N	6	12	11	14	9
ROLEPR	M	100.0	50.8	55.4	65.4	65.5
ROLEPR	S	0.0	29.6	17.9	13.7	16.4
ROLEPR	N	6	12	11	14	9
IFLTPR	M	68.0	40.7	51.3	48.6	43.8
IFLTPR	S	11.5	28.0	13.3	11.3	18.9
IFLTPR	N	6	12	11	14	9
IFWDPR	M	24.8	12.0	3.6	20.7	7.4
IFWDPR	S	4.9	19.2	3.8	16.3	8.0
IFWDPR	N	6	12	11	14	9
IFSTPR	M	7.3	37.3	31.8	18.9	24.8
IFSTPR	S	6.8	30.0	21.1	4.9	22.6
IFSTPR	N	6	12	11	14	9
MAT1PR	M	76.1	94.8	99.2	95.8	91.3
MAT1PR	S	13.0	5.9	1.0	6.1	6.2
MAT1PR	N	6	12	11	14	[illegible]
MAT2PR	M	52.7	79.1	76.5	73.1	65.7
MAT2PR	S	22.0	12.8	12.8	13.3	15.5
MAT2PR	N	6	12	11	14	9
ATINPR	M	25.2	53.9	56.2	39.2	41.0
ATINPR	S	24.7	22.0	18.1	16.2	20.5
ATINPR	N	6	12	11	14	9
ATLRPR	M	68.2	37.9	30.6	44.4	44.8
ATLRPR	S	23.6	22.7	18.7	18.9	19.2
ATLRPR	N	6	12	11	14	9
IFTTPR	M	100.0	90.0	86.7	[illegible]	75.9
IFTTPR	S	.0	11.1	18.8	10.8	18.2
IFTTPR	N	6	12	11	14	9
IFT2PR	M	32.1	49.3	35.4	39.5	32.2
IFT2PR	S	11.5	24.6	20.8	17.1	22.6
IFT2PR	N	6	12	11	14	9
ATTTPR	M	93.3	91.8	86.8	83.6	85.8
ATTTPR	S	5.0	5.9	16.9	7.8	7.4
ATTTPR	N	6	12	11	14	9
IFLTRP	M	67.9	44.1	61.7	56.5	60.1
IFLTRP	S	11.5	28.6	18.4	15.4	28.0
IFLTRP	N	6	12	11	14	9
IFWDRP	M	81.7	28.4	12.5	42.2	33.4
IFWDRP	S	15.1	38.2	12.4	22.4	30.9
IFWDRP	N	6	12	11	14	8
ATLRRP	M	73.9	41.1	[illegible]	53.0	52.9
ATLRRP	S	24.4	25.2	20.5	20.5	23.0
ATLRRP	N	6	12	11	14	9
STTMRT	M	94.6	100.0	100.0	95.1	89.6
STTMRT	S	9.8	0.0	0.0	10.1	18.8
STTMRT	N	6	12	11	14	9

Table 5

Reading and Mathematics Observation System - English:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 2 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
NSTRMN	M	12.3	21.4	15.9	16.2	16.6
NSTRMN	S	3.0	1.9	2.9	3.4	5.8
NSTRMN	N	13	28	34	30	17
CLSSMN	M	5.9	6.3	6.3	7.0	7.3
CLSSMN	S	0.3	0.5	0.3	0.1	0.0
CLSSMN	N	13	28	34	30	17
ROLEMN	M	3.9	3.3	4.3	5.3	6.2
ROLEMN	S	0.5	1.3	0.8	1.2	1.4
ROLEMN	N	13	28	34	30	17
RSOCMN	M	6.3	7.3	8.2	8.2	7.3
RSOCMN	S	1.7	0.3	0.3	0.6	0.9
RSOCMN	N	13	28	34	30	17
IFLTMN	M	1.1	1.8	2.0	1.9	2.0
IFLTMN	S	0.2	0.3	0.1	0.2	0.1
IFLTMN	N	13	26	34	30	16
IFWOMN	M	2.8	2.7	2.5	2.5	2.8
IFWOMN	S	0.4	0.2	0.3	0.7	0.4
IFWOMN	N	11	24	16	29	15
IFSTMN	M	1.6	[illegible]	1.4	2.0	1.4
IFSTMN	S	0.5	0.2	0.4	0.4	0.4
IFSTMN	N	8	23	34	29	17
TECHMN	M	1.1	1.3	1.7	1.8	1.8
TECHMN	S	0.2	0.2	0.2	0.2	0.3
TECHMN	N	13	27	34	30	17
LANGMN	M	2.5	2.8	2.9	2.9	2.9
LANGMN	S	0.3	0.3	0.2	0.2	0.2
LANGMN	N	13	28	34	30	17
MAT1MN	M	5.0	4.9	6.6	6.6	6.3
MAT1MN	S	1.0	0.7	0.9	0.7	1.0
MAT1MN	N	13	28	34	30	17
MAT2MN	M	3.0	3.3	3.9	4.7	4.5
MAT2MN	S	1.7	0.7	0.6	0.6	0.9
MAT2MN	N	13	25	34	30	17
ATNTMN	M	1.9	2.0	1.9	1.9	1.8
ATNTMN	S	0.1	.0	0.1	0.1	0.1
ATNTMN	N	13	28	34	30	17
ATINMN	M	1.1	1.1	1.6	1.9	2.0
ATINMN	S	0.2	0.3	0.3	0.4	0.2
ATINMN	N	13	25	34	30	15
ATLRMN	M	1.7	1.7	2.0	2.3	1.9
ATLRMN	S	0.5	0.4	0.5	0.3	0.6
ATLRMN		13	28	34	29	17

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
NEX2MN	M		2.1	2.9	1.6	1.7
NEX2MN	S		1.2	2.0	1.0	1.1
NEX2MN	N		17	34	30	17
PROCMN	M	3.1	3.2	3.3	3.1	3.1
PROCMN	S	0.4	0.3	0.5	0.5	0.3
PROCMN	N	13	28	34	30	17
NOISMN	M	2.9	2.3	2.1	2.0	2.2
NOISMN	S	0.2	0.3	0.4	0.4	0.4
NOISMN	N	13	28	34	30	17
ATT1MN	M	3.0	2.5			
ATT1MN	S	0.4	1.2			
ATT1MN	N	13	11			
CTRLMN	M	0.4	0.8	0.8	0.4	0.2
CTRLMN	S	0.4	1.4	0.6	0.6	0.2
CTRLMN	N	13	23	34	30	17
OBTMMN	M	21.5	31.9	22.2	24.1	23.4
OBTMMN	S	2.1	1.2	2.3	4.9	4.4
OBTMMN	N	13	25	33	30	17
STTMMN	M	21.5	31.6	31.9	24.1	23.4
STTMMN	S	2.1	2.4	1.7	4.9	4.4
STTMMN	N	13	28	34	30	17

BEST COPY AVAILABLE

Table 5

Reading and Mathematics Observation System - English:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for Site 2 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
CLSSPR	M	95.5	95.9	57.7	62.0	71.8
CLSSPR	S	9.1	13.1	28.8	25.8	21.4
CLSSPR	N	13	28	34	30	17
ROLEPR	M	95.5	95.9	65.1	64.8	80.8
ROLEPR	S	9.1	13.1	23.3	23.4	14.3
ROLEPR	N	13	28	34	30	17
IFLTPR	M	34.6	57.5	55.7	46.8	29.4
IFLTPR	S	19.8	23.5	12.4	19.7	22.2
IFLTPR	N	13	28	34	30	17
IFWDPR	M	23.3	9.7	6.1	9.4	16.9
IFWDPR	S	21.0	8.0	7.4	7.3	11.8
IFWDPR	N	13	28	34	30	17
IFSTPR	M	6.4	19.2	31.9	32.7	39.2
IFSTPR	S	7.3	12.7	20.4	21.9	24.5
IFSTPR	N	13	28	34	30	17
MAT1PR	M	87.4	97.4	96.4	94.3	92.4
MAT1PR	S	12.5	2.5	6.3	5.6	7.5
MAT1PR	N	13	28	34	30	17
MAT2PR	M	44.7	54.6	70.8	66.4	66.6
MAT2PR	S	15.4	27.2	21.3	18.1	28.9
MAT2PR	N	13	25	34	30	17
ATINPR	M	34.8	22.3	44.5	37.6	27.7
ATINPR	S	18.7	16.4	18.0	20.8	21.4
ATINPR	N	13	28	34	30	17
ATLRPR	M	50.0	71.9	46.8	51.0	52.8
ATLRPR	S	12.0	17.0	20.9	25.1	25.5
ATLRPR	N	13	28	34	30	17
IFTTPR	M	64.3	86.4	93.8	88.9	85.5
IFTTPR	S	17.2	8.8	7.1	8.2	10.5
IFTTPR	N	13	28	34	30	17
IFT2PR	M	29.7	28.9	38.0	42.1	56.1
IFT2PR	S	21.4	[illegible]	18.1	20.4	21.1
IFT2PR	N	13	28	34	30	17
ATTTPR	M	84.8	94.3	91.3	88.5	80.1
ATTTPR	S	[illegible]	3.1	9.9	7.8	14.5
ATTTPR	N	13	28	34	30	17
IFLTRP	M	49.2	64.7	59.2	52.7	34.0
IFLTRP	S	32.7	23.6	19.0	22.1	23.5
IFLTRP	N	13	28	34	30	17
IFWDRP	M	70.1	36.5	17.9	29.5	28.4
IFWDRP	S	31.5	22.0	22.0	27.3	18.9
IFWDRP	N	12	25	34	30	17
ATLRRP	M	58.5	76.4	50.1	56.7	64.0
ATLRRP	S	16.4	16.7	20.3	25.0	27.8
ATLRRP	N	13	28	34	30	17
STTMRT	M	100.0	99.1	99.1	100.0	100.0
STTMRT	S	0.0	4.9	1.9	0.0	0.0
STTMRT	N	13	28	34	30	17

BEST COPY AVAILABLE

Table 7

Reading and Mathematics Observation System - English:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 3 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
MSTDMN	M	7.3	12.5	17.9		
MSTDMN	S	2.2	2.7	2.7		
MSTDMN	N	73	71	76		
CLSSMN	M	5.8	5.6	6.4		
CLSSMN	S	0.4	0.4	0.5		
CLSSMN	N	73	71	76		
ROLEMN	M	7.2	5.2	5.9		
ROLEMN	S	0.8	0.5	0.5		
ROLEMN	N	73	71	76		
OBJCMN	M	8.1	8.2	8.3		
OBJCMN	S	0.5	0.7	0.7		
OBJCMN	N	73	71	76		
CFLTMN	M	1.5	2.0	2.1		
CFLTMN	S	0.4	0.2	0.1		
CFLTMN	N	74	71	74		
CFNDMN	M	3.0	2.0	2.3		
CFNDMN	S	0.0	0.0	0.5		
CFNDMN	N	17	2	50		
CFSTMN	M	1.9	1.5	1.6		
CFSTMN	S	0.1	0.3	0.3		
CFSTMN	N	21	71	73		
TECHMN	M	1.0	1.0	1.0		
TECHMN	S	0.1	0.0	0.0		
TECHMN	N	73	71	75		
LANGMN	M	3.0	2.9	3.0		
LANGMN	S	0.0	0.2	0.1		
LANGMN	N	73	71	76		
MAT1MN	M	4.4	5.7	6.0		
MAT1MN	S	2.2	0.4	0.5		
MAT1MN	N	73	71	76		
MAT2MN	M	4.1	4.3	5.0		
MAT2MN	S	1.6	0.8	1.0		
MAT2MN	N	28	71	73		
ATNTMN	M	1.9	2.0	2.0		
ATNTMN	S	0.2	.0	.0		
ATNTMN	N	73	71	76		
ATINMN	M	1.7	1.4	1.9		
ATINMN	S	0.4	0.2	0.2		
ATINMN	N	10	71	76		
ATLRMN	M	2.3	2.3	2.7		
ATLRMN	S	0.4	0.3	0.2		
ATLRMN	N	73	71	76		
MEN2MN	M	0.3	0.8	1.1		
MEN2MN	S	0.4	0.4	1.1		
MEN2MN	N	73	71	76		
PROCMN	M	2.9	2.8	2.7		
PROCMN	S	0.2	0.1	0.1		
PROCMN	N	72	71	76		
NOISMN	M	2.3	2.1	2.2		
NOISMN	S	0.3	0.1	0.2		
NOISMN	N	73	[illegible]	75		
ATT1MN	M					
ATT1MN	S					
ATT1MN	N					
CTRLMN	M	0.0	0.1	0.2		
CTRLMN	S	0.0	0.1	0.3		
CTRLMN	N	73	71	75		
OBTMMN	M	41.6	44.5	55.6		
OBTMMN	S	3.2	[illegible]	5.4		
OBTMMN	N	73	71	75		
STTMMN	M	39.0	44.4	53.6		
STTMMN	S	10.5	5.2	5.6		
STTMMN	N	73	71	76		

Table 3

BEST COPY AVAILABLE

Reading and Mathematics Observation System - English:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for Site 3 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
CLSSPR	M	89.1	60.6	74.7		
CLSSPR	S	16.1	12.4	11.1		
CLSSPR	N	35	71	76		
ROLEPR	M	89.1	61.1	74.7		
ROLEPR	S	16.1	12.0	11.1		
ROLEPR	N	35	71	76		
IFLTPR	M	75.3	48.0	66.5		
IFLTPR	S	20.3	13.9	15.5		
IFLTPR	N	35	71	76		
IFWDPR	M	7.0	0.1	4.1		
IFWDPR	S	10.5	0.5	9.2		
IFWDPR	N	35	71	76		
IFSTPR	M	11.8	46.5	23.6		
IFSTPR	S	12.4	12.1	12.4		
IFSTPR	N	35	71	76		
MAT1PR	M	93.0	97.2	96.2		
MAT1PR	S	5.9	2.4	6.7		
MAT1PR	N	35	71	76		
MAT2PR	M	48.0	62.5	42.8		
MAT2PR	S	26.3	13.8	18.1		
MAT2PR	N	28	71	73		
ATINPR	M	10.4	70.4	51.5		
ATINPR	S	19.1	10.4	14.1		
ATINPR	N	35	71	76		
ATLRPR	M	80.9	26.2	45.6		
ATLRPR	S	28.4	9.4	13.9		
ATLRPR	N	35	71	76		
IFTTPR	M	94.1	94.6	94.3		
IFTTPR	S	8.6	5.1	4.4		
IFTTPR	N	35	71	76		
IFT2PR	M	18.8	46.6	27.8		
IFT2PR	S	17.4	12.1	17.0		
IFT2PR	N	35	71	76		
ATTTPR	M	91.3	96.7	97 2		
ATTTPR	S	23.0	2.6	2.9		
ATTTPR	N	35	71	76		
IFLTRP	M	79.2	50.	70.9		
IFLTRP	S	21.6	13.6	.7.1		
IFLTRP	N	35	71	76		
IFWDRP	M	36.0	0.1	13.3		
IFWDRP	S	24.6	0.9	20.3		
IFWDRP	N	22	71	75		
ATLRRP	M	88.5	27.3	47.0		
ATLRRP	S	20.4	10.1	14.2		
ATLRRP	N	33	71	76		
STTMRT	M	94.8	99.9	94.9		
STTMRT	S	17.4	0.5	6.4		
STTMRT	N	35	71	76		

Table 9

Reading and Mathematics Observation System - English:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 5 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
NSTDMN	M	15.9	14.7	15.8		
NSTDMN	S	4.8	4.4	4.9		
NSTDMN	N	56	51	53		
CLSSMN	M	6.5	6.5	7.0		
CLSSMN	S	0.6	1.1	.0		
CLSSMN	N	56	48	51		
ROLEMN	M	5.9	5.0	6.6		
ROLEMN	S	1.3	1.1	1.5		
ROLEMN	N	56	48	51		
SBJCMN	M	5.6	7.6	8.3		
SBJCMN	S	1.8	0.6	0.4		
SBJCMN	N	56	51	53		
IFLTMN	M	1.8	2.1	2.0		
IFLTMN	S	0.6	0.3	0.2		
IFLTMN	N	55	50	51		
IFWDMN	M	1.0	2.9	2.5		
IFWDMN	S	0.0	0.2	0.7		
IFWDMN	N	7	20	42		
IFSTMN	M	2.2	1.2	1.1		
IFSTMN	S	0.8	0.3	0.2		
IFSTMN	N	20	42	46		
TECHMN	M	1.0	1.1	1.0		
TECHMN	S	0.0	0.3	0.0		
TECHMN	N	40	14	53		
LANGMN	M	2.8	2.8	3.0		
LANGMN	S	0.2	0.3	0.1		
LANGMN	N	56	51	53		
MAT1MN	M	3.4	5.5	5.7		
MAT1MN	S	1.1	0.9	0.6		
MAT1MN	N	56	51	53		
MAT2MN	M	3.2	4.3	4.4		
MAT2MN	S	1.5	0.9	0.7		
MAT2MN	N	42	49	52		
ATNTMN	M	1.9	1.9	2.0		
ATNTMN	S	0.1	0.1	0.1		
ATNTMN	N	56	51	53		
ATINMN	M	1.6	1.5	1.6		
ATINMN	S	0.6	0.4	0.5		
ATINMN	N	44	51	53		
ATLRMN	M	1.8	2.4	2.3		
ATLRMN	S	0.4	0.3	0.4		
ATLRMN	N	56	48	48		
NEX2MN	M	2.1	2.2	- 0.2		
NEX2MN	S	1.7	1.9	0.6		
NEX2MN	N	56	51	53		
PROCMN	M	3.2	3.0	3.0		
PROCMN	S	0.3	0.1	0.2		
PROCMN	N	55	50	52		
NOISMN	M	2.4	2.2	2.1		
NOISMN	S	0.3	0.2	0.2		
NOISMN	N	56	51	53		
ATTIMN	M					
ATTIMN	S					
ATTIMN	N					
CTRLMN	-	0.2	0.5	0.5		
CTRLMN	S	0.2	0.8	0.5		
CTRLMN	N	56	51	53		
OBTMMN	M	58.7	46.7	38.7		
OBTMMN	S	5.3	6.0	3.8		
OBTMMN	N	56	51	53		
STTMMN	M	55.6	44.0	36.9		
STTMMN	S	5.9	7.6	6.0		
STTMMN	N	56	51	53		

Table 10

Reading and Mathematics Observation System - English:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for Site 5 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
CLSSPR	M	65.3	56.0	56.8		
CLSSPR	S	19.2	25.5	27.3		
CLSSPR	N	56	51	53		
ROLEPR	M	66.3	65.2	62.8		
ROLEPR	S	18.7	21.1	25.6		
ROLEPR	N	56	48	51		
IFLTPR	M	34.9	50.7	47.5		
IFLTPR	S	22.6	25.1	21.5		
IFLTPR	N	56	51	53		
IFWDPR	M	0.3	2.2	11.1		
IFWDPR	S	0.7	3.6	9.7		
IFWDPR	N	56	51	53		
IFSTPR	M	3.4	19.9	33.3		
IFSTPR	S	5.3	23.0	24.6		
IFSTPR	N	56	51	53		
MAT1PR	M	82.5	89.9	93.5		
MAT1PR	S	10.5	8.0	6.9		
MAT1PR	N	56	51	53		
MAT2PR	M	30.5	64.4	70.8		
MAT2PR	S	22.7	18.2	17.6		
MAT2PR	N	42	49	52		
ATINPR	M	23.9	53.9	56.8		
ATINPR	S	17.2	22.8	20.9		
ATINPR	N	56	51	53		
ATLRPR	M	60.1	34.6	37.4		
ATLRPR	S	20.2	21.4	20.5		
ATLRPR	N	56	51	53		
IFTTPR	M	38.6	72.8	91.9		
IFTTPR	S	20.8	19.7	10.8		
IFTTPR	N	56	51	53		
IFT2PR	M	3.7	22.1	44.4		
IFT2PR	S	5.7	22.2	22.9		
IFT2PR	N	56	51	53		
ATTTPR	M	84.0	88.5	94.1		
ATTTPR	S	9.2	7.9	5.7		
ATTTPR	N	[illegible]	51	53		
IFLTRP	M	88.0	69.5	52.3		
IFLTRP	S	18.9	26.1	23.6		
IFLTRP	N	55	51	53		
IFWDRP	M	5.7	17.3	33.5		
IFWDRP	S	8.0	26.4	26.9		
IFWDRP	N	20	42	50		
ATLRRP	M	71.0	39.6	39.8		
ATLRRP	S	20.9	23.7	21.9		
ATLRRP	N	56	51	53		
STTMRT	M	95.1	94.8	94.9		
STTMRT	S	8.6	11.9	12.0		
STTMRT	N	56	51	53		

APPENDIX F

Reading and Mathematics Observation System - Spanish:

Descriptive Statistics on the Instructional Indices by Instructional Year for Each Site for the Bilingual Sample

Table 1

Reading and Mathematics Observation System - Spanish:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 0 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
NSTDMN	M	10.6	13.2	10.0	19.0	19.3
NSTDMN	S	1.4	3.0	1.2	7.4	3.7
NSTDMN	N	20	9	15	12	7
CLSSMN	M	3.1	5.6	3.6	5.5	5.3
CLSSMN	S	0.1	0.6	0.4	0.6	0.6
CLSSMN	N	20	9	15	12	7
ROLEMN	M	5.1	5.5	7.4	7.2	7.7
ROLEMN	S	1.0	1.1	0.3	1.1	0.4
ROLEMN	N	20	9	15	12	7
OBJCMN	M	3.1	5.3	5.0	3.2	7.9
OBJCMN	S	0.9	0.4	0.4	0.5	0.4
OBJCMN	N	20	9	15	12	7
IFLTMN	M	1.9	2.2	2.4	2.5	2.1
IFLTMN	S	0.7	0.2	0.2	0.4	0.2
IFLTMN	N	19	9	15	11	7
IFWDMN	M	3.0	3.0	3.0	3.0	3.0
IFWDMN	S	0.0	0.0	0.1	0.0	0.0
IFWDMN	N	12	7	12	10	9
IFSTMN	M	2.4	2.0	1.7	2.2	1.6
IFSTMN	S	0.3	0.3	0.5	0.5	0.6
IFSTMN	N	12	9	14	10	7
TECHMN	M	1.3	1.2	1.5	1.4	1.4
TECHMN	S	0.3	0.2	0.2	0.4	0.2
TECHMN	N	20	9	15	12	7
LANGMN	M	1.4	1.0	1.1	1.2	1.3
LANGMN	S	0.1	0.0	0.1	0.3	0.1
LANGMN	N	20	9	15	12	7
MAT1MN	M	3.6	5.6	6.1	6.8	3.0
MAT1MN	S	0.6	1.1	1.0	0.8	0.0
MAT1MN	N	20	9	15	12	7
MAT2MN	M	4.1	4.6	3.6	3.5	3.0
MAT2MN	S	0.8	1.1	0.7	0.8	0.1
MAT2MN	N	20	9	15	12	7
ATNTMN	M	1.9	1.8	1.9	1.9	1.8
ATNTMN	S	0.1	.0	.0	0.1	0.1
ATNTMN	N	20	9	15	12	7
ATINMN	M	1.0	1.3	1.9	1.7	
ATINMN	S	0.0	0.4	0.1	0.4	
ATINMN	N	5	9	15	9	
ATLRMN	M	2.3	1.8	2.3	2.4	2.3
ATLRMN	S	0.2	0.3	0.2	0.4	0.3
ATLRMN	N	20	9	15	12	7

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
NENZMN	M			1.3	3.6	3.7
NENZMN	S			0.6	1.6	2.2
NENZMN	N			9	12	7
PROCMN	M	3.3	3.8	3.7	3.3	3.3
PROCMN	S	0.4	0.3	0.2	0.3	0.1
PROCMN	N	20	9	15	12	7
NOISMN	M	2.7	2.6	2.6	2.3	2.9
NOISMN	S	0.4	0.3	0.2	0.3	0.4
NOISMN	N	20	9	15	12	7
ATTIMN	M	3.7	3.8	4.0		
ATTIMN	S	0.3	0.3	.0		
ATTIMN	N	20	9	9		
CTRLMN	M	0.9	2.1	0.0	0.0	0.0
CTRLMN	S	0.5	1.8	0.0	0.0	0.0
CTRLMN	N	20	9	15	12	7
OBTMMN	M	50.7	63.0	57.0	42.3	32.3
OBTMMN	S	3.3	0.0	2.4	3.9	3.4
OBTMMN	N	20	9	15	12	7
STTMMN	M	49.9	39.2	34.8	42.3	32.3
STTMMN	S	3.3	12.1	3.2	3.9	3.4
STTMMN	N	20	9	15	12	7

Table 2

Reading and Mathematics Observation System - Spanish:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for Site 0 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
CLSSPR	M	87.6	94.0	95.1	90.5	100.0
CLSSPR	S	13.5	7.1	4.1	12.7	0.0
CLSSPR	N	20	9	15	12	7
ROLEPR	M	87.6	94.0	95.1	90.5	100.0
ROLEPR	S	13.5	7.1	4.1	12.7	0.0
ROLEPR	N	20	9	15	12	7
IFLTPR	M	55.1	46.3	49.2	31.8	31.3
IFLTPR	S	34.6	14.3	13.9	27.0	21.5
IFLTPR	N	20	9	15	12	7
IFWDPR	M	18.2	4.6	23.2	23.4	24.9
IFWDPR	S	18.8	5.9	19.3	20.7	11.0
IFWDPR	N	20	9	15	12	7
IFSTPR	M	19.5	15.6	14.8	32.9	26.9
IFSTPR	S	17.9	6.9	13.8	30.4	4.3
IFSTPR	N	20	9	15	12	7
MAT1PR	M	95.7	81.0	89.0	93.9	85.8
MAT1PR	S	2.7	5.9	4.5	5.7	5.1
MAT1PR	N	20	9	15	12	7
MAT2PR	M	78.3	32.2	68.6	78.8	83.0
MAT2PR	S	14.2	15.0	9.8	16.7	6.4
MAT2PR	N	20	9	15	12	7
ATINPR	M	8.5	13.9	16.3	24.1	0.0
ATINPR	S	15.4	5.8	11.0	21.6	0.0
ATINPR	N	20	9	15	12	7
ATLRPR	M	84.8	65.4	70.0	69.1	85.8
ATLRPR	S	15.0	9.7	9.6	22.8	5.1
ATLRPR	N	20	9	15	12	7
IFTTPR	M	92.3	66.5	87.1	88.1	83.0
IFTTPR	S	8.0	5.2	4.1	6.0	6.4
IFTTPR	N	20	9	15	12	7
IFT2PR	M	37.7	20.2	37.9	56.4	51.8
IFT2PR	S	34.6	13.6	16.1	22.6	15.2
IFT2PR	N	20	9	15	12	7
ATTTPR	M	93.2	79.3	86.3	93.1	85.8
ATTTPR	S	3.2	4.2	4.4	5.1	5.1
ATTTPR	N	20	9	15	12	7
IFLTRP	M	58.1	69.7	57.6	34.8	36.3
IFLTRP	S	36.6	20.9	17.9	27.5	20.2
IFLTRP	N	20	9	15	12	7
IFWDRP	M	44.9	12.1	65.3	42.7	43.3
IFWDRP	S	18.8	18.2	31.3	40.7	19.1
IFWDRP	N	12	9	14	12	7
ATLRRP	M	91.0	82.1	81.5	74.0	100.0
ATLRRP	S	16.4	8.7	12.4	23.4	0.0
ATLRRP	N	20	9	15	12	7
STTMRT	M	97.8	62.3	66.0	100.0	100.0
STTMRT	S	4.9	19.2	15.9	0.0	0.0
STTMRT	N	20	9	15	12	7

Table 3

Reading and Mathematics Observation System - Spanish:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 1 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
NSTDMN	M	12.3	15.8	16.1	20.1	
NSTDMN	S	2.2	1.6	5.2	5.9	
NSTDMN	N	7	8	9	5	
CLSSMN	M	5.5	6.6	6.1	6.3	
CLSSMN	S	0.5	0.3	0.4	0.3	
CLSSMN	N	7	8	9	5	
ROLEMN	M	5.5	4.6	5.2	4.2	
ROLEMN	S	1.2	1.3	0.7	1.1	
ROLEMN	N	7	9	8	5	
SBJCMN	M	6.6	3.1	7.9	3.0	
SBJCMN	S	1.5	0.2	0.5	0.9	
SBJCMN	N	7	8	8	5	
IFLTMN	M	1.0	1.7	1.3	1.9	
IFLTMN	S	.0	0.3	0.3	0.2	
IFLTMN	N	7	8	8	4	
IFWOMN	M	3.0	2.0	2.8	1.8	
IFWOMN	S	0.0	0.0	0.4	0.8	
IFWOMN	N	7	2	5	5	
IFSTMN	M	2.0	1.4	1.4	1.5	
IFSTMN	S	0.0	0.3	0.4	0.7	
IFSTMN	N	4	3	9	3	
TECHMN	M	1.2	1.1	1.3	1.6	
TECHMN	S	0.1	0.1	0.4	0.3	
TECHMN	N	7	8	8	5	
LANGMN	M	1.7	1.0	1.1	1.2	
LANGMN	S	0.1	.0	0.1	0.2	
LANGMN	N	7	8	9	5	
MAT1MN	M	5.2	5.4	5.5	5.5	
MAT1MN	S	0.7	0.5	0.6	0.7	
MAT1MN	N	7	9	8	5	
MAT2MN	M	1.3	5.1	4.6	4.8	
MAT2MN	S	0.9	0.2	0.6	0.4	
MAT2MN	N	7	9	8	5	
ATNTMN	M	2.0	2.0	1.9	2.0	
ATNTMN	S	.0	0.1	0.1	0.1	
ATNTMN	N	7	8	8	5	
ATINMN	M	1.0	1.4	1.1	2.0	
ATINMN	S	.	0.3	0.2	0.5	
ATINMN	N	7	8	8	5	
ATLRMN	M	1.2	1.7	1.8	2.0	
ATLRMN	S	0.4	0.4	0.6	0.6	
ATLRMN	N	7	3	9	5	

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
MEN2MN	M		1.0	3.9	3.0	
MEN2MN	S		0.4	3.4	1.9	
MEN2MN	N		4	8	5	
PROCMN	M	2.9	3.0	2.9	3.2	
PROCMN	S	0.3	0.1	0.5	0.2	
PROCMN	N	7	9	3	5	
NOISMN	M	2.7	2.5	2.4	2.0	
NOISMN	S	0.5	0.2	0.4	0.2	
NOISMN	N	7	3	9	5	
ATT1MN	M	3.1	3.2			
ATT1MN	S	0.1	0.2			
ATT1MN	N	7	4			
CTRLMN	M	1.0	0.6	0.2	0.1	
CTRLMN	S	0.8	0.3	0.3	0.2	
CTRLMN	N	7	9	6	5	
OBTMMN	M	31.9	31.6	30.4	30.7	
OBTMMN	S	3.2	0.8	0.8	0.7	
OBTMMN	N	7	8	3	5	
STTMMN	M	31.9	31.6	30.0	30.7	
STTMMN	S	3.2	0.8	1.2	0.7	
STTMMN	N	7	9	9	5	

770

Table 4

Reading and Mathematics Observation System - Spanish:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for Site 1 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
CLSSPR	M	93.8	38.9	73.9	59.0	
CLSSPR	S	9.2	19.0	17.3	25.5	
CLSSPR	N	7	8	8	5	
ROLEPR	M	93.8	43.4	74.8	61.5	
ROLEPR	S	9.2	11.8	16.0	24.6	
ROLEPR	N	7	8	8	5	
IFLTPR	M	48.0	53.3	56.6	43.9	
IFLTPR	S	17.0	13.5	18.2	30.0	
IFLTPR	N	7	8	8	5	
IFWDPR	M	22.9	0.4	7.5	29.3	
IFWDPR	S	14.5	0.8	11.5	9.4	
IFWDPR	N	7	8	8	5	
IFSTPR	M	6.3	44.6	32.7	18.5	
IFSTPR	S	7.2	14.6	17.8	23.2	
IFSTPR	N	7	8	8	5	
MAT1PR	M	95.1	98.4	98.7	100.0	
MAT1PR	S	4.9	2.2	3.1	0.0	
MAT1PR	N	7	8	8	5	
MAT2PR	M	71.9	71.1	71.5	98.6	
MAT2PR	S	9.1	15.6	17.8	3.1	
MAT2PR	N	7	8	8	5	
ATINPR	M	49.3	57.7	42.9	69.4	
ATINPR	S	12.5	16.1	10.5	20.8	
ATINPR	N	7	8	8	5	
ATLRPR	M	50.0	36.9	46.2	26.8	
ATLRPR	S	12.4	14.1	17.4	19.4	
ATLRPR	N	7	8	8	5	
IFTTPR	M	77.1	98.3	96.7	91.6	
IFTTPR	S	18.9	2.4	5.0	18.8	
IFTTPR	N	7	8	8	5	
IFT2PR	M	29.1	45.0	40.2	47.7	
IFT2PR	S	12.0	15.1	19.1	30.1	
IFT2PR	N	7	8	8	5	
ATTTPR	M	99.3	94.6	89.1	96.1	
ATTTPR	S	1.9	7.8	12.0	4.1	
ATTTPR	N	7	8	8	5	
IFLTRP	M	62.1	54.4	58.8	48.4	
IFLTRP	S	12.3	14.5	19.1	28.6	
IFLTRP	N	7	8	8	5	
IFWDRP	M	72.5	0.9	15.4	71.2	
IFWDRP	S	33.0	1.7	21.4	26.8	
IFWDRP	N	7	8	8	5	
ATLRRP	M	50.6	38.3	49.8	27.1	
ATLRRP	S	12.4	15.4	19.1	19.5	
ATLRRP	N	7	8	8	5	
STTMRT	M	100.0	100.0	98.7	100.0	
STTMRT	S	0.0	0.0	3.7	0.0	
STTMRT	N	7	8	8	5	

Table 5

Reading and Mathematics Observation System - Spanish:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 2 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
NSTDMN	M	11.6	14.7	17.6	22.4	24.0
NSTDMN	S	4.0	5.6	5.3	2.8	0.0
NSTDMN	N	13	15	12	3	2
CLSSMN	M	3.7	5.0	5.9	7.0	7.0
CLSSMN	S	0.4	0.7	0.2	0.0	0.0
CLSSMN	N	13	15	12	3	2
ROLEMN	M	5.7	5.9	4.0	5.3	5.5
ROLEMN	S	1.3	0.5	1.2	0.3	0.0
ROLEMN	N	13	15	12	3	2
SBJCMN	M	5.3	3.0	5.7	3.9	9.0
SBJCMN	S	1.7	0.6	1.3	0.1	0.0
SBJCMN	N	13	15	12	3	2
IFLTMN	M	1.1	1.7	2.0	2.0	2.0
IFLTMN	S	0.1	0.3	0.0	0.0	0.0
IFLTMN	N	13	15	12	3	2
IFWGMN	M	2.8	2.9			
IFWGMN	S	0.3	0.2			
IFWGMN	N	13	14			
IFSTMN	M	1.5	1.7	1.1	1.5	1.5
IFSTMN	S	0.5	0.3	0.1	0.0	0.0
IFSTMN	N	8	15	12	2	2
TECHMN	M	1.1	1.5	1.6	2.0	2.0
TECHMN	S	0.1	0.2	0.3	0.0	0.0
TECHMN	N	13	15	12	3	2
LANGMN	M	1.4	1.5	1.3	1.7	1.6
LANGMN	S	0.2	0.1	0.2	0.1	0.0
LANGMN	N	13	15	12	3	2
MAT1MN	M	8.0	4.4	6.2	5.7	4.9
MAT1MN	S	0.9	0.3	0.8	1.3	0.0
MAT1MN	N	13	15	12	3	2
MAT2MN	M	4.6	4.7	4.5	8.0	8.0
MAT2MN	S	0.6	0.4	0.5	0.0	0.0
MAT2MN	N	12	15	10	2	2
ATNTMN	M	1.9	2.0	1.9	1.8	1.3
ATNTMN	S	0.1	0.1	0.1	0.1	0.0
ATNTMN	N	13	15	12	3	2
ATINMN	M	1.1	1.1	1.6	2.0	
ATINMN	S	0.1	0.1	0.3	0.0	
ATINMN	N	13	15	12	1	
ATLRMN	M	1.4	1.8	2.0	2.1	2.6
ATLRMN	S	0.3	0.2	0.2	0.9	0.0
ATLRMN	N	13	15	12	3	2

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
MEN2MN	M		10.6	3.8	3.9	11.7
MEN2MN	S		1.3	1.7	4.9	0.0
MEN2MN	N		5	12	3	2
PROCMN	M	3.0	2.6	2.7	2.3	2.3
PROCMN	S	0.4	0.4	0.4	0.1	0.0
PROCMN	N	13	15	12	3	2
NOISMN	M	2.9	2.7	2.8	2.7	3.0
NOISMN	S	0.3	0.5	0.5	0.5	0.0
NOISMN	N	13	15	12	3	2
ATT1MN	M	3.2	3.3			
ATT1MN	S	0.2	0.1			
ATT1MN	N	13	10			
CTRLMN	M	0.4	0.3	0.0	1.3	1.0
CTRLMN	S	0.4	0.4	0.0	0.6	0.0
CTRLMN	N	13	15	12	3	2
OBTMMN	M	30.7	30.9	31.0	27.7	25.0
OBTMMN	S	2.3	1.2	0.9	2.9	0.0
OBTMMN	N	13	15	12	3	2
STTMMN	M	30.1	30.3	31.0	27.7	25.0
STTMMN	S	3.4	1.0	1.0	2.9	0.0
STTMMN	N	13	15	12	3	2

Table 6

Reading and Mathematics Observation System - Spanish:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for Site 2 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
CLSSPR	M	86.9	77.5	90.2	83.9	100.0
CLSSPR	S	19.5	11.9	19.6	27.9	0.0
CLSSPR	N	13	15	12	3	2
ROLEPR	M	86.9	78.0	90.2	83.9	100.0
ROLEPR	S	19.5	11.0	19.6	27.9	0.0
ROLEPR	N	13	15	12	3	2
IFLTPR	M	56.0	21.4	40.5	56.4	34.6
IFLTPR	S	25.6	15.6	20.5	37.8	0.0
IFLTPR	N	13	7	12	3	2
IFWDPR	M	19.8	5.3	0.0	0.0	0.0
IFWDPR	S	17.1	2.2	0.0	0.0	0.0
IFWDPR	N	13	15	12	3	2
IFSTPR	M	3.6	23.5	38.5	30.8	46.2
IFSTPR	S	3.4	10.6	10.9	26.7	0.0
IFSTPR	N	13	15	12	3	2
MAT1PR	M	93.5	99.2	75.9	97.2	80.8
MAT1PR	S	5.8	1.1	16.1	11.1	0.0
MAT1PR	N	13	15	12	3	2
MAT2PR	M	49.0	56.2	32.2	50.0	50.0
MAT2PR	S	16.7	10.7	16.1	0.0	0.0
MAT2PR	N	12	15	10	2	2
ATINPR	M	26.6	33.3	45.9	6.5	0.0
ATINPR	S	20.0	11.7	26.3	11.2	0.0
ATINPR	N	13	15	12	3	2
ATLRPR	M	61.5	62.5	44.5	78.6	80.8
ATLRPR	S	15.0	9.6	22.0	3.8	0.0
ATLRPR	N	13	15	12	3	2
IFTTPR	M	79.4	88.1	79.1	87.2	80.8
IFTTPR	S	11.6	11.1	14.3	11.1	0.0
IFTTPR	N	13	15	12	3	2
IFT2PR	M	23.4	28.8	38.5	30.8	46.2
IFT2PR	S	18.5	12.3	10.9	26.7	0.0
IFT2PR	N	13	15	12	3	2
ATTTPR	M	88.0	95.8	90.4	85.1	80.8
ATTTPR	S	8.5	5.1	9.9	7.4	0.0
ATTTPR	N	13	15	12	3	2
IFLTRP	M	66.8	65.7	46.8	61.9	42.8
IFLTRP	S	26.5	16.2	19.2	33.0	0.0
IFLTRP	N	13	15	12	3	2
IFWDRP	M	81.9	16.5	0.0	0.0	0.0
IFWDRP	S	21.8	8.0	0.0	0.0	0.0
IFWDRP	N	13	15	12	2	2
ATLRRP	M	70.7	65.6	48.9	93.1	100.0
ATLRRP	S	19.8	11.4	24.5	12.0	0.0
ATLRRP	N	13	15	12	3	2
STTMRT	M	97.6	98.0	100.0	100.0	100.0
STTMRT	S	6.0	2.3	0.0	0.0	0.0
STTMRT	N	13	15	12	3	2

BEST COPY AVAILABLE

Table 7

Reading and Mathematics Observation System - Spanish:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 3 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
NSTDMN	M		9.3	9.6		
NSTDMN	S		3.1	0.9		
NSTDMN	N		7	5		
CLSSMN	M		6.4	6.5		
CLSSMN	S		[illegible]	0.2		
CLSSMN	N		6	5		
ROLEMN	M		5.9	7.1		
ROLEMN	S		0.4	0.9		
ROLEMN	N		6	5		
SBJCMN	M		3.7	3.3		
SBJCMN	S		0.4	0.0		
SBJCMN	N		7	5		
IFLTMN	M		2.1	2.2		
IFLTMN	S		0.1	0.2		
IFLTMN	N		6	5		
IFNOMN	M					
IFNOMN	S					
IFNOMN	N					
IFSTMN	M		1.2	1.6		
IFSTMN	S		0.3	0.1		
IFSTMN	N		7	5		
TECHMN	M		1.0	1.0		
TECHMN	S		0.0	0.0		
TECHMN	N		7	5		
LANGMN	M		1.1	1.2		
LANGMN	S		0.1	0.3		
LANGMN	N		7	5		
MAT1MN	M		4.9	6.3		
MAT1MN	S		0.9	0.3		
MAT1MN	N		7	5		
MAT2MN	M		4.0	4.6		
MAT2MN	S		0.8	2.1		
MAT2MN	N		7	5		
ATNTMN	M		2.0	2.0		
ATNTMN	S		0.0	0.0		
ATNTMN	N		7	5		
ATINMN	M		1.1	2.1		
ATINMN	S		0.1	0.3		
ATINMN	N		7	5		
ATLRMN	M		2.1	2.7		
ATLRMN	S		0.3	0.2		
ATLRMN	N		7	5		

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
MEN2MN	M		0.3	1.3		
MEN2MN	S		0.2	0.3		
MEN2MN	N		7	5		
PROCMN	M		2.9	2.7		
PROCMN	S		0.1	0.1		
PROCMN	N		7	5		
NOISMN	M		2.1	2.2		
NOISMN	S		0.1	0.1		
NOISMN	N		7	5		
ATT1MN	M					
ATT1MN	S					
ATT1MN	N					
CTRLMN	M		0.3	0.0		
CTRLMN	S		0.2	0.0		
CTRLMN	N		7	5		
OBTMMN	M		42.7	52.3		
OBTMMN	S		9.0	0.7		
OBTMMN	N		7	5		
STTMMN	M		42.7	50.6		
STTMMN	S		9.0	0.2		
STTMMN	N		7	5		

Table 8

Reading and Mathematics Observation System - Spanish: Descriptive Statistics on the Quantity Indices of Instruction by Instructional Year for Site 3 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
CLSSPR	M		47.6	62.8		
CLSSPR	S		27.7	19.4		
CLSSPR	N		7	5		
ROLEPR	M		55.5	62.8		
ROLEPR	S		19.8	19.4		
ROLEPR	N		6	5		
IFLTPR	M		21.4	51.2		
IFLTPR	S		15.6	13.8		
IFLTPR	N		7	5		
IFWDPR	M		0.0	0.0		
IFWDPR	S		0.0	0.0		
IFWDPR	N		7	5		
IFSTPR	M		75.2	43.9		
IFSTPR	S		15.0	13.9		
IFSTPR	N		7	5		
MAT1PR	M		99.6	95.1		
MAT1PR	S		0.3	0.7		
MAT1PR	N		7	5		
MAT2PR	M		59.8	25.5		
MAT2PR	S		24.3	10.9		
MAT2PR	N		7	5		
ATINPR	M		67.2	41.7		
ATINPR	S		23.7	21.0		
ATINPR	N		7	5		
ATLRPR	M		32.0	57.6		
ATLRPR	S		24.3	21.0		
ATLRPR	N		7	5		
IFTTPR	M		96.6	95.0		
IFTTPR	S		4.8	0.7		
IFTTPR	N		7	5		
IFT2PR	M		75.2	43.9		
IFT2PR	S		15.0	13.9		
IFT2PR	N		7	5		
ATTTPR	M		99.2	99.4		
ATTTPR	S		0.8	0.3		
ATTTPR	N		7	5		
IFLTRP	M		22.0	53.8		
IFLTRP	S		15.7	14.4		
IFLTRP	N		7	5		
IFWDRP	M		0.0	0.0		
IFWDRP	S		0.0	0.0		
IFWDRP	N		7	5		
ATLRRP	M		32.1	58.0		
ATLRRP	S		24.2	21.1		
ATLRRP	N		7	5		
STTMRT	M		100.0	97.3		
STTMRT	S		0.0	1.5		
STTMRT	N		7	5		

BEST COPY AVAILABLE

Table 9

Reading and Mathematics Observation System - Spanish:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 5 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
ASTDMN	M	13.0	11.4	14.[illegible]		
ASTDMN	S	5.1	3.3	7.7		
ASTDMN	N	27	34	22		
CLSSMN	M	5.6	5.6	7.0		
CLSSMN	S	0.5	1.2	0.0		
CLSSMN	N	27	31	22		
ROLEMN	M	5.7	5.1	7.3		
ROLEMN	S	1.4	1.4	1.0		
ROLEMN	N	27	31	22		
OBJCMN	M	5.8	7.7	7.9		
OBJCMN	S	1.7	0.7	1.4		
OBJCMN	N	27	34	22		
IFLTMN	M	2.0	2.1	2.1		
IFLTMN	S	0.5	0.4	0.3		
IFLTMN	N	27	31	21		
IFWDMN	M		3.0	2.7		
IFWDMN	S		0.0	0.5		
IFWDMN	N		9	5		
IFSTMN	M	1.0	1.3	1.3		
IFSTMN	S	0.0	0.4	0.3		
IFSTMN	N	19	20	15		
TECHMN	M	1.0	1.0	1.0		
TECHMN	S	0.0	0.0	0.0		
TECHMN	N	20	7	3		
LANGMN	M	1.6	1.2	1.3		
LANGMN	S	0.5	0.3	0.4		
LANGMN	N	27	34	22		
MAT1MN	M	4.5	5.6	5.9		
MAT1MN	S	1.5	0.8	1.1		
MAT1MN	N	27	34	22		
MAT2MN	M	2.7	3.9	3.9		
MAT2MN	S	1.4	0.9	1.0		
MAT2MN	N	23	34	22		
ATNTMN	M	1.9	1.9	2.0		
ATNTMN	S	0.1	0.1	.0		
ATNTMN	N	27	34	22		
ATINMN	M	1.3	1.5	1.8		
ATINMN	S	0.5	0.4	0.6		
ATINMN	N	21	34	20		
ATLRMN	M	2.0	2.2	2.3		
ATLRMN	S	0.6	0.4	0.4		
ATLRMN	N	27	32	22		

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
MENZMN	M	2.0	2.0	.0		
MENZMN	S	1.5	.9	.0		
MENZMN	N	27	34	22		
PROCMN	M	3.2	3.0	2.9		
PROCMN	S	0.3	0.1	0.2		
PROCMN	N	27	34	21		
NOISMN	M	2.2	2.2	2.1		
NOISMN	S	0.3	0.2	0.2		
NOISMN	N	27	34.0	22		
ATTIMN	M					
ATTIMN	S					
ATTIMN	N					
CTRLMN	M	0.1	0.5	0.3		
CTRLMN	S	0.3	0.5	0.4		
CTRLMN	N	27	34	22		
OBTMMN	M	52.4	49.4	73.3		
OBTMMN	S	14.2	4.0	11.9		
OBTMMN	N	27	34	22		
STTMMN	M	50.3	48.5	74.0		
STTMMN	S	14.3	4.5	11.9		
STTMMN	N	27	34	22		

776

Table 10

Reading and Mathematics Observation System - Spanish:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for Site 5 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
CLSSPR	M	81.6	43.0	82.7		
CLSSPR	S	22.2	24.5	23.4		
CLSSPR	N	27	34	22		
ROLEPR	M	81.6	51.9	86.5		
ROLEPR	S	22.2	19.9	21.6		
ROLEPR	N	27	31	22		
IFLTPR	M	50.3	54.6	64.8		
IFLTPR	S	17.5	21.7	29.4		
IFLTPR	N	27	34	22		
IFWDPR	M	0.0	0.9	6.8		
IFWDPR	S	0.0	2.3	13.4		
IFWDPR	N	27	34	22		
IFSTPR	M	1.6	24.0	21.3		
IFSTPR	S	2.5	21.5	23.5		
IFSTPR	N	27	34	22		
MAT1PR	M	82.1	92.1	99.2		
MAT1PR	S	13.8	7.6	1.5		
MAT1PR	N	27	34	22		
MAT2PR	M	46.9	65.0	68.2		
MAT2PR	S	23.8	17.0	28.0		
MAT2PR	N	23	34	22		
ATINPR	M	17.8	58.1	43.4		
ATINPR	S	17.9	22.7	32.7		
ATINPR	N	27	34	22		
ATLRPR	M	66.3	31.8	55.1		
ATLRPR	S	19.8	22.0	32.7		
ATLRPR	N	27	34	22		
IFTTPR	M	51.9	79.5	92.8		
IFTTPR	S	17.7	10.5	20.6		
IFTTPR	N	27	34	22		
IFT2PR	M	1.6	24.9	28.1		
IFT2PR	S	2.5	21.4	23.4		
IFT2PR	N	27	34	22		
ATTTPR	M	84.2	89.9	98.5		
ATTTPR	S	6.8	7.1	4.3		
ATTTPR	N	27	34	22		
IFLTRP	M	96.9	68.3	65.5		
IFLTRP	S	5.5	27.4	28.9		
IFLTRP	N	27	34	22		
IFWDRP	M	0.0	7.8	28.7		
IFWDRP	S	0.0	16.8	40.5		
IFWDRP	N	10	30	16		
ATLRRP	M	78.5	35.9	55.9		
ATLRRP	S	21.5	23.5	32.9		
ATLRRP	N	27	34	22		
STTMRT	M	96.6	98.3	97.7		
STTMRT	S	8.9	5.5	6.2		
STTMRT	N	27	34	22		

APPENDIX G

Teacher Checklists:

Teacher Instructional Plan Form
Master Code Sheet

BEST COPY AVAILABLE

TEACHER INSTRUCTIONAL PLAN

Teacher ____________________

School ____________________

Grade ____________________

Period __________ to __________
(date) (date)

READING GROUPS:

Group Name or I.D. No.	No. of Students	INITIALS OF STUDENTS IN GROUP																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

GROUP/ STUDENTS	Strategy Number	Instructional Focus	MATERIALS: Type	MATERIALS: Title	MATERIALS: Page(s)	Activity	Language	Instructor	Role	Total Minutes/ Two-week Period	SPECIAL NOTES

SEDL READING PROJECT

Checklist Master Code Sheet

Instructional Focus

S General skills

SAL Alphabetizing
SAD Auditory discrimination
SCA Composing aloud
SCL Cap-lower case corres.
SDU Dictionary usage
SDW Dictating word
SLN Letter-name work
SOR Oral reading fluency
SPD Phonetic discrimination
SSA Skimming ability
SSL Syllabification
SSP Spelling practice
SSR Silent reading fluency
SVD Visual discrimination
SWL Writing letters
SWP Writing phrases
SWS Writing sentences
SWW Writing words

G Grammar skills

GAD Adjective, adverb, etc.
GCN Contractions
GCP Capitalization
GCW Compound words
GND Noun derivative
GPU Punctuation usage
GSP Singular/plural
GSS Sentence structure
GTU Tense usage
GVD Verb derivation

V Vocabulary skills

VAS Antonyms/synonyms
VEC Easily confusable words
VVE Vocabulary enrichment

D Decoding strategies

DCR Clause recognition
DCS Letter cluster - sound recognition
DLR Letter recognition
DLS Letter-sound recognition
DME Morpheme element recognition
DPR Phrase recognition
DSP Spelling pattern recognition
DSR Simple sentence recognition
DWT Whole text recognition
DWW Whole word recognition

C Comprehension strategies

CLF Literal facts
CMD Major ideas
CMN Making inferences
CRE Relations
CSS Story sequence
CTA Talking about what read

N Interpretation strategies

NAH Appreciating humor
NEA Emotional attitude
NPE Predicting events
NSM Sensory images
NSR Seeing relationships

Materials (Type)

LC Letter card
PC Phonics card
PH Phrase card
PR Picture card
SC Sentence card
TC Task card
WC Word card

Materials (Type) cont'd

AW Another workbook
BW Basal workbook
DS Ditto sheet
PT Programmed text
PW Phonics workbook

BT Basal text
CT Commercial test
ST Standardized test
TT Teacher-made test
NF Informal Reading Inventory

BR Basal reader
LB Library Book
SB Supplementary book

AM Art material
AS Auditory stimulus
AV Audio-visual aid
CB Chalkboard
CC Commercial chart
CH Chart - teacher-made
GA Game
LM Language master
MN Magazine, newspaper
PP Paper, pencil
QT Questions from teacher
SM Student-made materials
SN (Other) stuff
TL Teacher manual
TM Teacher-made materials
TW Typewriter
TR Tape recorder

Activity

A__ Act'y reading related
_AA Art activity
_DA Dramatization activity
_EW Evaluating work
_MA Music activity
_MC Making correspondences
_PA Picture activity
_PG Playing games
_RN Reading instructions
_RV Reciting/responding verse

L__ Lecture
_LL Listen lecture
_LS Listen story

D__ Discussion
_SP Speak

W__ Independent work
_CM Copying material
_CR Creating writing
_LD Listen, do
_RS Reading silently
_TA Teaching another
_WA Writing answers
_WP Writing practice

Q__ Question, answers
_SA Spoken answers
_TT Test taking
_WA Written answers
_WD Writing from dictation

R__ Recitation, reading
_CA Composing aloud
_DR Drill
_DS Dictating story
_RA Reading aloud
_RR Reading and responding
_RU Reading in unison

Activity (cont'd)

V__ Audio-visual
_WL Watch, listen

T__ Transitional activity
_CU Clean-up
_CW Completing work
_PW Preparing work
_WT Wait time

N__ Not reading

Language of Instruction

S Spanish
E English
A Alternate use of both
C Code switching

Instructor

A Aide
C Cross-age tutor
D Administrator
N Intern teacher
P Peer tutor
R Resource teacher
S Substitute teacher
T Teacher
U Other adult not usually in the room
V Volunteer

Role

A Assess/diagnose
C Control
D Discuss
F Facilitate
L Lecture
M Manage class
N Instruction
P Prepare
R Read
S Show
V Observing

781

APPENDIX H

Checklists - English:

Descriptive Statistics on the Instructional Indices by Instructional Year for Each Site for the Bilingual Sample

Table 1

Checklists - English:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 0 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTMN	M	1.7	2.1	2.2	2.3	2.0
IFLTMN	S	0.7	0.4	0.2	0.2	0.1
IFLTMN	N	32	53	54	51	34
IFWDMN	M	2.9	2.8	2.7	2.7	2.9
IFWDMN	S	0.2	0.2	0.4	0.6	0.3
IFWDMN	N	32	31	45	43	33
IFSTMN	M	2.5	1.6	1.6	2.0	1.8
IFSTMN	S	0.6	0.6	0.6	0.6	0.6
IFSTMN	N	32	46	54	49	34
MATIMN	M	4.5	6.3	6.9	7.4	7.5
MATIMN	S	1.3	1.6	0.9	0.5	0.3
MATIMN	N	32	53	54	51	34
NBRDMN	M	1.5	1.2	1.5	1.2	1.0
NBRDMN	S	0.5	0.6	0.4	0.5	0.1
NBRDMN	N	32	53	54	51	34
ATINMN	M	1.8	2.0	2.1	2.1	2.2
ATINMN	S	0.5	0.4	0.2	0.2	0.2
ATINMN	N	26	42	53	50	34
ATLRMN	M	1.9	2.2	2.4	2.5	2.4
ATLRMN	S	0.5	0.4	0.2	0.2	0.2
ATLRMN	N	32	53	54	51	34
CLSSMN	M	6.6	6.8	6.9	6.9	6.6
CLSSMN	S	0.4	0.3	0.1	0.2	0.7
CLSSMN	N	32	53	54	51	34
ROLEMN	M	8.7	8.5	8.3	7.7	7.9
ROLEMN	S	0.3	0.4	0.4	0.4	0.4
ROLEMN	N	32	53	54	51	34
NSTDMN	M	8.3	9.2	8.2	9.9	12.5
NSTDMN	S	2.1	2.6	2.4	3.1	5.0
NSTDMN	N	32	53	54	51	34
TIMEMN	M	440.8	343.9	376.8	543.2	579.3
TIMEMN	S	134.5	155.5	97.5	141.3	132.2
TIMEMN	N	32	53	54	51	34
RNKIMN	M		2.0	1.8	2.2	2.2
RNKIMN	S		0.9	0.8	0.8	0.9
RNKIMN	N		39	40	42	21
RMKEMN	M		1.4	1.4	1.7	1.6
RMKEMN	S		0.5	0.5	0.7	0.5
RMKEMN	N		44	54	51	34
NCISMN	M	4.4	3.3	2.5	1.5	1.5
NCISMN	S	0.5	1.7	1.5	0.2	0.2
NCISMN	N	32	53	54	51	34

Table 2

Checklists - English:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for Site 0 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTPR	M	43.6	73.8	67.2	50.5	39.1
IFLTPR	S	13.4	11.9	13.6	16.0	15.3
IFLTPR	N	32	53	54	51	34
IFWDPR	M	36.4	8.8	9.3	13.5	17.5
IFWDPR	S	15.9	14.7	10.5	9.1	13.9
IFWDPR	N	32	53	54	51	34
IFSTPR	M	20.0	17.4	23.5	36.0	43.4
IFSTPR	S	6.0	9.3	11.2	15.6	18.1
IFSTPR	N	32	53	54	51	34
ATINPR	M	16.7	13.6	20.7	35.9	31.2
ATINPR	S	10.6	13.2	12.3	11.1	10.7
ATINPR	N	32	53	54	51	34
ATLRPR	M	83.2	86.4	79.2	64.1	68.7
ATLRPR	S	10.6	13.2	12.3	11.1	10.7
ATLRPR	N	32	53	54	51	34
IFT2PR	M	56.3	26.2	32.8	49.5	61.0
IFT2PR	S	13.4	11.9	13.6	16.0	15.3
IFT2PR	N	32	53	54	51	34
IFWDRP	M	60.5	24.6	26.3	26.8	30.7
IFWDRP	S	18.2	31.9	24.0	19.4	21.2
IFWDRP	N	32	51	54	50	34

784

BEST COPY AVAILABLE

Table 3

Checklists - English:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 1 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTMN	M	2.5	2.1	2.2	2.1	2.1
IFLTMN	S	0.0	0.0	0.2	0.1	0.1
IFLTMN	N	3	9	13	14	9
IFWDMN	M		3.0	2.5	2.3	2.3
IFWDMN	S		.0	0.5	0.4	0.5
IFWDMN	N		6	8	14	9
IFSTMN	M	2.5	1.4	1.4	1.7	1.8
IFSTMN	S	0.0	0.4	0.4	0.4	0.3
IFSTMN	N	3	9	11	13	9
MAT1MN	M	5.2	6.5	6.8	7.6	7.2
MAT1MN	S	0.0	0.7	0.5	0.4	0.4
MAT1MN	N	3	9	13	14	9
NBRDMN	M	1.0	1.0	1.0	0.9	0.9
NBRDMN	S	0.0	0.0	0.0	0.3	0.3
NBRDMN	N	3	9	13	14	9
ATINMN	M	1.3	1.4	1.9	2.2	2.0
ATINMN	S	0.0	0.4	0.5	0.2	0.0
ATINMN	N	3	9	10	14	9
ATLRMN	M	2.0	2.4	2.3	2.4	2.4
ATLRMN	S	0.0	0.2	0.3	0.2	0.1
ATLRMN	N	3	9	13	14	9
CLSSMN	M	7.0	7.0	7.0	7.0	7.0
CLSSMN	S	0.0	0.0	0.0	0.0	0.0
CLSSMN	N	3	9	13	14	9
ROLEMN	M	7.3	8.0	8.6	7.8	7.5
ROLEMN	S	0.0	0.5	0.3	0.4	0.4
ROLEMN	N	3	9	13	14	9
NSTDMN	M	13.0	7.3	9.9	11.8	11.8
NSTDMN	S	0.0	0.8	2.5	3.1	4.4
NSTDMN	N	3	9	13	14	9
TIMEMN	M	246.7	522.8	359.3	718.6	653.3
TIMEMN	S	0.0	149.8	46.9	190.2	191.7
TIMEMN	N	3	9	13	14	9
RNKIMN	M			2.4	2.1	2.0
RNKIMN	S			0.7	0.7	1.0
RNKIMN	N			6	11	7
RMKEMN	M		1.0	1.4	1.3	1.5
RMKEMN	S		0.1	0.7	0.4	0.5
RMKEMN	N		9	13	14	9
NCISMN	M	9.0	6.3	1.3	1.5	1.6
NCISMN	S	0.0	4.8	0.3	0.2	0.3
NCISMN	N	3	9	13	14	9

785

Table 4

Checklists - English:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for Site 1 Bilingual Sample

Scale	Statistic	IY		IY2	IY3	IY4
IFLTPR	M	86.9	84.1	79.2	66.3	57.7
IFLTPR	S	0.0	11.2	14.5	7.9	16.8
IFLTPR	N	3	9	13	14	9
IFWDPR	M	0.0	4.1	8.1	22.2	17.9
IFWDPR	S	0.0	3.3	11.4	10.9	12.6
IFWDPR	N	3	9	13	14	9
IFSTPR	M	13.0	11.8	12.6	11.5	24.3
IFSTPR	S	0.0	8.2	10.4	7.7	14.9
IFSTPR	N	3	9	13	14	9
ATINPR	M	30.8	30.4	21.7	35.3	32.5
ATINPR	S	0.0	9.3	16.6	8.8	9.1
ATINPR	N	3	9	13	14	9
ATLRPR	M	69.1	69.5	78.3	64.7	67.5
ATLRPR	S	0.0	9.3	16.6	8.8	9.1
ATLRPR	N	3	9	[illegible]	14	9
IFT2PR	M	13.0	15.9	20.8	33.7	42.3
IFT2PR	S	0.0	11.2	14.5	7.9	16.8
IFT2PR	N	3	9	13	14	9
IFWDRP	M	0.0	17.6	23.2	56.1	36.9
IFWDRP	S	0.0	14.3	30.0	28.4	19.9
IFWDRP	N	3	9	12	14	9

780

Table 5

Checklists - English:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 2 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTMN	M	1.4	2.2	2.1	2.2	2.0
IFLTMN	S	0.4	0.1	0.1	0.1	0.1
IFLTMN	N	12	21	35	32	17
IFWDMN	M	3.0	2.0	2.3	2.1	1.5
IFWDMN	S	0.0	0.0	0.4	0.6	0.5
IFWDMN	N	7	1	16	31	17
IFSTMN	M	1.3	2.6	2.2	2.2	2.0
IFSTMN	S	0.4	0.5	0.2	0.2	0.3
IFSTMN	N	9	31	25	31	17
MAT1MN	M	6.2	5.6	6.9	7.4	7.5
MAT1MN	S	1.4	2.1	0.3	0.3	0.2
MAT1MN	N	12	31	35	32	17
NBRDMN	M	1.0	0.7	1.1	1.0	1.0
NBRDMN	S	0.0	0.5	0.1	0.0	0.0
NBRDMN	N	12	31	35	32	17
ATINMN	M	2.7	1.9	2.0	2.3	2.0
ATINMN	S	0.7	0.1	0.3	0.4	.0
ATINMN	N	12	31	34	30	17
ATLRMN	M	2.1	2.2	2.2	2.5	2.3
ATLRMN	S	0.1	0.5	0.7	0.2	0.1
ATLRMN	N	12	27	35	32	17
CLSSMN	M	7.0	7.0	7.0	7.0	7.0
CLSSMN	S	0.0	0.0	0.0	0.0	0.0
CLSSMN	N	12	31	35	32	17
ROLEMN	M	8.2	7.1	7.4	8.2	7.5
ROLEMN	S	0.1	0.4	0.5	0.6	0.6
ROLEMN	N	12	31	35	32	17
NSTDMN	M	6.0	6.9	8.0	14.5	18.5
NSTDMN	S	0.0	2.3	2.6	5.5	4.8
NSTDMN	N	12	31	35	32	17
TIMEMN	M	150.4	356.7	497.1	748.5	565.9
TIMEMN	S	1.4	252.2	100.7	131.4	229.2
TIMEMN	N	12	31	35	32	17
RNKIMN	M	2.6	1.8	2.1	2.2	2.0
RNKIMN	S	0.5	0.6	0.7	0.9	1.4
RNKIMN	N	5	21	35	17	2
RMKEMN	M	1.6	1.6	1.6	1.8	2.0
RMKEMN	S	0.5	0.5	0.5	0.3	0.1
RMKEMN	N	5	21	35	32	17
NCISMN	M	4.6	4.8	1.4	1.3	1.5
NCISMN	S	3.1	3.1	0.2	0.2	0.1
NCISMN	N	12	31	35	32	17

Table 6

Checklists - English:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for Site 2 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTPR	M	79.3	45.9	72.8	63.4	48.2
IFLTPR	S	15.2	33.4	21.1	13.6	10.3
IFLTPR	N	12	31	35	32	17
IFWDPR	M	5.9	0.1	7.4	14.6	25.1
IFWDPR	S	5.7	0.3	9.7	8.5	12.8
IFWDPR	N	12	31	35	32	17
IFSTPR	M	14.7	54.1	19.8	21.9	26.7
IFSTPR	S	12.3	33.5	16.5	7.3	10.7
IFSTPR	N	12	31	35	32	17
ATINPR	M	19.0	52.7	44.6	15.3	34.5
ATINPR	S	9.3	27.6	13.0	10.1	12.5
ATINPR	N	12	31	35	32	17
ATLRPR	M	81.0	47.3	55.4	84.7	65.5
ATLRPR	S	9.2	27.6	13.0	10.1	12.5
ATLRPR	N	12	31	35	32	17
IFT2PR	M	20.7	54.1	27.2	36.5	51.8
IFT2PR	S	15.2	33.4	21.1	13.6	10.3
IFT2PR	N	12	31	35	32	17
IFWDRP	M	31.0	0.3	23.1	40.2	39.8
IFWDRP	S	21.2	1.7	21.1	14.8	17.2
IFWDRP	N	9	31	25	31	17

788

Table 7

Checklists - English:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 3 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTMN	M	1.6	2.1	2.2		
IFLTMN	S	0.6	0.3	0.2		
IFLTMN	N	60	70	75		
IFWDMN	M	3.0	3.0	2.7		
IFWDMN	S	0.0	0.2	0.4		
IFWDMN	N	47	41	60		
IFSTMN	M	2.4	2.2	2.3		
IFSTMN	S	0.7	0.5	0.5		
IFSTMN	N	57	66	71		
MATIMN	M	4.1	5.2	7.1		
MATIMN	S	1.2	1.1	0.4		
MATIMN	N	60	70	75		
NBRDMN	M	0.2	0.6	1.0		
NBRDMN	S	0.3	0.3	0.1		
NBRDMN	N	60	70	75		
ATINMN	M	1.3	2.1	2.0		
ATINMN	S	0.3	0.5	0.2		
ATINMN	N	59	67	75		
ATLRMN	M	2.3	2.2	2.3		
ATLRMN	S	0.3	0.3	0.2		
ATLRMN	N	60	70	75		
CLSSMN	M	6.7	6.9	7.0		
CLSSMN	S	0.3	0.2	0.1		
CLSSMN	N	60	70	75		
ROLEMN	M	7.9	8.2	7.7		
ROLEMN	S	0.3	0.4	0.5		
ROLEMN	N	60	70	75		
NSTDMN	M	13.4	8.4	9.3		
NSTDMN	S	6.5	2.2	3.1		
NSTDMN	N	60	70	75		
TIMEMN	M	440.8	465.6	686.2		
TIMEMN	S	147.9	177.2	19[illegible].1		
TIMEMN	N	60	70	75		
RNKIMN	M	2.5	2.0	2.1		
RNKIMN	S	0.6	0.8	0.8		
RNKIMN	N	4	32	53		
RMKEMN	M	2.8	1.2	1.3		
RMKEMN	S	0.4	0.4	0.5		
RMKEMN	N	16	68	75		
NCISMN	M	1.0	1.4	1.4		
NCISMN	S	0.0	0.2	0.2		
NCISMN	N	60	70	75		

Table 8

Checklists - English:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for Site 3 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTPR	M	54.9	78.1	70.0		
IFLTPR	S	21.5	16.0	9.3		
IFLTPR	N	60	70	75		
IFWDPR	M	20.5	5.1	9.0		
IFWDPR	S	14.8	6.1	7.0		
IFWDPR	N	60	70	75		
IFSTPR	M	24.6	16.7	21.0		
IFSTPR	S	17.5	12.7	10.4		
IFSTPR	N	60	70	75		
ATINPR	M	27.7	28.6	28.5		
ATINPR	S	14.1	16.9	11.7		
ATINPR	N	60	70	75		
ATLRPR	M	72.3	71.3	71.5		
ATLRPR	S	14.1	16.9	11.7		
ATLRPR	N	60	70	75		
IFT2PR	M	45.1	21.9	30.0		
IFT2PR	S	21.5	16.0	9.3		
IFT2PR	N	60	70	75		
IFWDRP	M	43.0	23.5	31.5		
IFWDRP	S	30.2	24.3	27.6		
IFWDRP	N	60	68	75		

730

Table 9

Checklists - English:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 5 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTMN	M	1.9	2.2	2.3		
IFLTMN	S	0.6	0.5	0.2		
IFLTMN	N	54	40	51		
IFWDMN	M	3.0	2.8	2.3		
IFWDMN	S	0.1	0.4	0.6		
IFWDMN	N	44	10	42		
IFSTMN	M	2.3	1.8	1.9		
IFSTMN	S	0.6	0.5	0.4		
IFSTMN	N	41	32	50		
MAT1MN	M	4.1	4.3	7.0		
MAT1MN	S	1.5	0.7	0.5		
MAT1MN	N	56	40	51		
NBRDMN	M	0.6	1.0	0.9		
NBRDMN	S	0.5	0.2	0.3		
NBRDMN	N	56	40	51		
ATINMN	M	1.6	2.2	2.1		
ATINMN	S	0.7	0.5	0.1		
ATINMN	N	44	38	50		
ATLRMN	M	2.4	2.3	2.4		
ATLRMN	S	0.3	0.3	0.3		
ATLRMN	N	56	36	51		
CLSSMN	M	6.4	7.0	7.0		
CLSSMN	S	0.7	0.1	0.0		
CLSSMN	N	56	40	51		
ROLEMN	M	8.3	7.9	7.6		
ROLEMN	S	0.4	0.6	0.5		
ROLEMN	N	56	40	51		
NSTDMN	M	9.0	8.4	9.9		
NSTDMN	S	6.0	2.5	4.5		
NSTDMN	N	56	40	51		
TIMEMN	M	308.2	346.2	832.4		
TIMEMN	S	69.6	97.9	178.1		
TIMEMN	N	56	40	51		
RNKIMN	M	2.9	2.4	2.0		
RNKIMN	S	0.4	0.7	0.6		
RNKIMN	N	7	14	45		
RMKEMN	M	2.0	1.6	1.7		
RMKEMN	S	0.1	0.4	0.5		
RMKEMN	N	44	38	51		
NCISMN	M	1.2	1.3	1.4		
NCISMN	S	0.3	0.3	0.2		
NCISMN	N	56	40	51		

731

Table 10

Checklists - English:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for Site 5 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTPR	M	62.2	79.2	69.4		
IFLTPR	S	29.3	14.9	10.5		
IFLTPR	N	56	40	51		
IFWDPR	M	22.6	3.9	12.0		
IFWDPR	S	23.7	7.7	10.6		
IFWDPR	N	56	40	51		
IFSTPR	M	15.3	16.9	18.7		
IFSTPR	S	20.3	11.9	7.5		
IFSTPR	N	56	40	51		
ATINPR	M	17.8	40.7	34.0		
ATINPR	S	14.8	26.1	11.5		
ATINPR	N	56	40	51		
ATLRPR	M	82.2	59.3	66.0		
ATLRPR	S	14.8	26.1	11.5		
ATLRPR	N	56	40	51		
IFT2PR	M	37.8	20.7	30.7		
IFT2PR	S	29.3	14.9	10.5		
IFT2PR	N	56	40	51		
IFWDRP	M	62.3	11.1	37.8		
IFWDRP	S	33.0	21.7	25.4		
IFWDRP	N	50	33	51		

732

APPENDIX I

Checklists - Spanish:

Descriptive Statistics on the Instructional Indices by Instructional Year for Each Site for the Bilingual Sample

Table 1

Checklists - Spanish:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 0 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTMN	M	2.9	2.4	2.3	2.2	2.1
IFLTMN	S	0.2	0.4	0.2	0.3	0.1
IFLTMN	N	26	17	38	35	18
IFWDMN	M	3.0		3.0	2.8	3.0
IFWDMN	S	0.0		0.1	0.2	0.0
IFWDMN	N	20		15	24	10
IFSTMN	M	1.9	2.1	1.5	1.1	1.8
IFSTMN	S	0.2	0.9	0.8	0.2	0.9
IFSTMN	N	9	15	24	19	18
MATIMN	M	4.0	6.9	7.1	7.5	7.1
MATIMN	S	1.2	0.7	0.7	0.5	1.5
MATIMN	N	26	17	38	35	19
NBRDMN	M	1.0	1.0	1.0	0.7	0.4
NBRDMN	S	0.2	0.0	0.1	0.6	0.4
NBRDMN	N	26	17	38	35	19
ATINMN	M	2.8	2.0	2.1	2.2	2.4
ATINMN	S	0.2	0.7	0.2	0.3	0.5
ATINMN	N	19	11	28	18	18
ATLRMN	M	2.2	1.7	2.4	2.2	2.5
ATLRMN	S	0.3	0.4	0.4	0.4	0.2
ATLRMN	N	26	17	38	35	19
CLSSMN	M	6.4	6.8	6.9	6.9	6.8
CLSSMN	S	0.5	0.3	0.2	0.2	0.4
CLSSMN	N	26	17	38	35	19
ROLEMN	M	8.9	8.6	8.2	8.5	8.7
ROLEMN	S	0.2	0.3	0.6	0.5	0.3
ROLEMN	N	26	17	38	35	19
NSTDMN	M	7.9	9.7	8.0	9.4	12.1
NSTDMN	S	2.1	2.6	2.5	2.4	5.9
NSTDMN	N	26	17	38	35	19
TIMEMN	M	189.4	212.6	183.5	259.9	350.3
TIMEMN	S	50.1	66.0	53.2	92.0	94.9
TIMEMN	N	26	17	38	35	19
RNKIMN	M		1.7	1.9	1.8	1.9
RNKIMN	S		0.6	0.3	0.4	0.2
RNKIMN	N		3	22	34	18
RMKEMN	M		1.0	1.2	1.2	1.3
RMKEMN	S		0.0	0.3	0.4	0.4
RMKEMN	N		3	24	34	18
NCISMN	M	4.5	1.7	2.0	1.5	1.3
NCISMN	S	2.3	0.4	1.0	0.5	0.3
NCISMN	N	26	17	38	35	19

Table 2

Checklists - Spanish:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for Site 0 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTPR	M	60.8	81.6	85.5	70.2	58.0
IFLTPR	S	20.8	12.2	13.2	22.5	19.8
IFLTPR	N	26	17	38	35	19
IFWDPR	M	36.5	0.0	6.0	18.7	15.5
IFWDPR	S	22.6	0.0	12.7	17.3	15.4
IFWDPR	N	26	17	38	35	19
IFSTPR	M	2.6	18.5	8.5	11.1	26.4
IFSTPR	S	3.9	12.2	9.0	16.1	18.9
IFSTPR	N	26	17	38	35	19
ATINPR	M	22.7	7.0	20.8	12.0	17.6
ATINPR	S	15.9	6.2	15.0	13.8	6.4
ATINPR	N	26	17	38	35	19
ATLRPR	M	77.3	93.0	79.2	88.0	82.4
ATLRPR	S	16.0	6.2	15.0	13.8	6.4
ATLRPR	N	26	17	38	35	19
IFT2PR	M	39.2	18.5	14.5	29.8	41.9
IFT2PR	S	20.8	12.2	13.2	22.5	19.8
IFT2PR	N	26	17	38	35	19
IFWDRP	M	71.2	0.0	32.4	57.2	32.6
IFWDRP	S	42.8	0.0	39.0	42.2	33.2
IFWDRP	N	26	15	29	33	19

795

Table 3

Checklists - Spanish:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 1 Bilingual Sample

Scale	Statistic	Site 1 IY0	IY1	IY2	IY3	IY4
IFLTMN	M	2.5	2.3	2.2	2.2	
IFLTMN	S	0.0	0.2	0.2	0.2	
IFLTMN	N	4	8	8	3	
IFWDMN	M		3.0	2.8	2.1	
IFWDMN	S		0.0	0.4	0.8	
IFWDMN	N		4	5	3	
IFSTMN	M	2.5	1.7	2.1	1.2	
IFSTMN	S	0.0	0.3	0.8	0.3	
IFSTMN	N	4	8	8	3	
MATIMN	M	5.0	5.3	6.7	7.1	
MATIMN	S	0.0	0.9	0.9	1.4	
MATIMN	N	4	8	8	3	
NBRDMN	M	0.0	0.8	0.9	0.8	
NBRDMN	S	0.0	0.3	0.2	0.3	
NBRDMN	N	4	8	8	3	
ATINMN	M	1.3	1.6	1.9	2.1	
ATINMN	S	0.0	0.6	0.4	0.1	
ATINMN	N	4	8	8	3	
ATLRMN	M	2.0	2.0	2.4	2.1	
ATLRMN	S	0.0	0.1	0.2	0.3	
ATLRMN	N	4	8	8	3	
CLSSMN	M	7.0	7.0	7.0	6.9	
CLSSMN	S	0.0	0.0	0.0	0.2	
CLSSMN	N	4	8	8	3	
ROLEMN	M	7.3	8.4	8.5	8.1	
ROLEMN	S	0.0	0.2	0.5	0.3	
ROLEMN	N	4	8	8	3	
NSTDMN	M	12.0	13.4	8.8	11.0	
NSTDMN	S	0.0	3.5	2.0	7.1	
NSTDMN	N	4	8	8	3	
TIMEMN	M	246.7	345.2	430.9	866.7	
TIMEMN	S	0.0	45.3	115.1	395.5	
TIMEMN	N	4	8	8	3	
RNKIMN	M		1.7	2.0		
RNKIMN	S		1.0	0.8		
RNKIMN	N		8	4		
RMKEMN	M		1.2	1.3	1.0	
RMKEMN	S		0.2	0.3	0.0	
RMKEMN	N		8	8	3	
NCISMN	M	9.0	5.3	1.6	1.7	
NCISMN	S	0.0	4.0	0.4	0.3	
NCISMN	N	4	8	8	3	

Table 4

Checklists - Spanish:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for Site 1 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTPR	M	86.9	71.5	72.2	64.9	
IFLTPR	S	0.0	9.9	16.3	10.5	
IFLTPR	N	4	8	8	3	
IFWDPR	M	0.0	5.0	14.3	23.9	
IFWDPR	S	0.0	5.3	14.0	8.5	
IFWDPR	N	4	8	8	3	
IFSTPR	M	13.0	23.5	13.5	11.3	
IFSTPR	S	0.0	8.5	5.4	5.8	
IFSTPR	N	4	8	8	3	
ATINPR	M	30.8	21.6	19.8	32.5	
ATINPR	S	0.0	4.1	11.2	12.0	
ATINPR	N	4	8	8	3	
ATLRPR	M	69.1	78.4	80.1	67.5	
ATLRPR	S	0.0	4.1	11.2	12.0	
ATLRPR	N	4	8	8	3	
IFT2PR	M	13.0	28.5	27.8	35.2	
IFT2PR	S	0.0	9.9	16.4	10.5	
IFT2PR	N	4	8	8	3	
IFWDRP	M	0.0	15.2	37.5	68.6	
IFWDRP	S	0.0	16.3	31.9	13.3	
IFWDRP	N	4	8	8	3	

737

Table 5

Checklists - Spanish:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 2 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTMN	M	1.4	2.2	2.0	2.2	2.2
IFLTMN	S	0.5	0.2	0.0	.0	0.0
IFLTMN	N	9	15	10	10	2
IFWDMN	M	3.0			3.0	
IFWDMN	S	0.0			0.0	
IFWDMN	N	3			8	
IFSTMN	M	2.0	2.6	1.5	1.5	2.4
IFSTMN	S	0.0	0.5	0.0	0.5	0.0
IFSTMN	N	4	9	10	10	2
MATIMN	M	3.1	6.7	6.2	8.0	7.9
MATIMN	S	1.1	0.5	[illegible]	0.1	0.0
MATIMN	N	9	15	10	10	2
NBRDMN	M	1.0	1.0	1.0	1.0	1.0
NBRDMN	S	0.0	0.0	0.2	0.0	0.0
NBRDMN	N	9	15	10	10	2
ATINMN	M	1.8	1.8	1.9	2.0	2.0
ATINMN	S	0.9	0.2	0.0	0.0	0.0
ATINMN	N	8	15	10	2	2
ATLRMN	M	1.7	2.5	2.4	2.5	1.8
ATLRMN	S	0.7	0.2	0.1	0.4	0.0
ATLRMN	N	9	15	10	10	2
CLSSMN	M	7.0	7.0	7.0	7.0	7.0
CLSSMN	S	0.0	0.0	0.0	0.0	0.0
CLSSMN	N	9	15	10	10	2
ROLEMN	M	8.2	8.2	7.9	8.9	8.4
ROLEMN	S	0.3	0.4	0.1	0.3	0.0
ROLEMN	N	9	15	10	10	2
NSTDMN	M	6.0	14.6	8.1	21.2	28.0
NSTDMN	S	0.0	8.5	1.2	3.6	0.0
NSTDMN	N	9	15	10	10	2
TIMEMN	M	150.6	409.7	328.8	267.0	255.0
TIMEMN	S	1.7	287.5	35.9	6.3	0.0
TIMEMN	N	9	15	10	10	2
RNKIMN	M		2.3	1.9		
RNKIMN	S		0.6	0.3		
RNKIMN	N		10	10		
RMKEMN	M		1.6	1.6	1.8	1.0
RMKEMN	S		0.5	0.3	0.4	0.0
RMKEMN	N		15	10	10	2
NCISMN	M	6.0	4.5	1.5	1.1	1.5
NCISMN	S	3.2	2.2	0.0	0.2	0.0
NCISMN	N	9	15	10	10	2

Table 6

Checklists - Spanish:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for Site 2 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTPR	M	94.3	95.5	74.0	61.2	80.9
IFLTPR	S	7.2	4.0	5.5	10.4	0.0
IFLTPR	N	9	15	10	10	2
IFWDPR	M	3.7	0.0	0.0	3.0	0.0
IFWDPR	S	5.6	0.0	0.0	4.2	0.0
IFWDPR	N	9	15	10	10	2
IFSTPR	M	2.0	4.4	25.9	30.8	19.1
IFSTPR	S	2.5	4.0	5.6	6.2	0.0
IFSTPR	N	9	15	10	10	2
ATINPR	M	11.5	22.8	40.2	1.9	9.6
ATINPR	S	6.2	11.7	5.6	4.0	0.0
ATINPR	N	9	15	10	10	2
ATLRPR	M	88.5	77.1	59.7	98.1	90.5
ATLRPR	S	6.2	11.7	5.5	4.0	0.0
ATLRPR	N	9	15	10	10	2
IFT2PR	M	5.7	4.4	25.9	38.8	19.1
IFT2PR	S	7.2	4.0	5.6	10.4	0.0
IFT2PR	N	9	15	10	10	2
IFWDRP	M	56.3	0.0	0.0	16.0	0.0
IFWDRP	S	37.5	0.0	0.0	8.4	0.0
IFWDRP	N	4	9	10	10	2

739

Table 7

Checklists - Spanish:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 3 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTMN	M	1.4	2.3	2.1		
IFLTMN	S	0.2	0.4	0.2		
IFLTMN	N	9	6	5		
IFWDMN	M	3.0	3.0	2.3		
IFWDMN	S	0.0	0.0	0.9		
IFWDMN	N	14	5	4		
IFSTMN	M	1.9	1.3	1.5		
IFSTMN	S	0.4	0.6	0.4		
IFSTMN	N	13	6	3		
MAT1MN	M	4.0	4.6	7.2		
MAT1MN	S	0.8	0.7	0.6		
MAT1MN	N	14	6	5		
NBRDMN	M	0.2	0.5	1.0		
NBRDMN	S	0.2	0.0	0.0		
NBRDMN	N	14	6	5		
ATINMN	M	1.0	1.3	2.0		
ATINMN	S	0.0	0.5	.0		
ATINMN	N	12	6	5		
ATLRMN	M	2.1	2.2	2.3		
ATLRMN	S	0.3	0.2	0.2		
ATLRMN	N	14	6	5		
CLSSMN	M	6.9	7.0	6.7		
CLSSMN	S	0.1	0.1	0.3		
CLSSMN	N	14	6	5		
ROLEMN	M	8.4	8.6	7.7		
ROLEMN	S	0.5	0.4	0.8		
ROLEMN	N	14	6	5		
NSTDMN	M	7.4	7.0	7.1		
NSTDMN	S	3.6	1.0	0.5		
NSTDMN	N	14	6	5		
TIMEMN	M	392.5	293.7	540.0		
TIMEMN	S	80.3	73.2	95.7		
TIMEMN	N	14	6	5		
RNKIMN	M	1.9	2.0	2.0		
RNKIMN	S	0.8	1.0	1.0		
RNKIMN	N	8	5	5		
RMKEMN	M	1.0	1.7	1.4		
RMKEMN	S	0.0	0.5	0.5		
RMKEMN	N	8	6	5		
NCISMN	M	1.0	1.5	1.1		
NCISMN	S	0.0	0.0	0.2		
NCISMN	N	14	6	5		

Table 8

Checklists - Spanish:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for Site 3 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTPR	M	37.1	70.0	76.9		
IFLTPR	S	28.8	12.0	21.5		
IFLTPR	N	14	6	5		
IFWDPR	M	29.7	17.5	7.0		
IFWDPR	S	31.0	11.7	4.2		
IFWDPR	N	14	6	5		
IFSTPR	M	33.3	12.5	16.1		
IFSTPR	S	15.4	10.6	18.7		
IFSTPR	N	14	6	5		
ATINPR	M	15.1	18.7	43.8		
ATINPR	S	8.8	15.4	20.0		
ATINPR	N	14	6	5		
ATLRPR	M	84.9	81.3	56.2		
ATLRPR	S	8.8	15.4	20.0		
ATLRPR	N	14	6	5		
IFT2PR	M	63.0	30.0	23.0		
IFT2PR	S	28.8	12.0	21.5		
IFT2PR	N	14	6	5		
IFWDRP	M	36.2	60.9	50.0		
IFWDRP	S	28.6	32.8	46.8		
IFWDRP	N	14	6	5		

8 11

Table 9

Checklists - Spanish:
Descriptive Statistics on the Quality Indices of Instruction
by Instructional Year for Site 5 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTMN	M	2.2	2.5	2.3		
IFLTMN	S	0.4	0.2	0.2		
IFLTMN	N	39	29	29		
IFWDMN	M	3.0	2.6	2.6		
IFWDMN	S	0.0	0.7	0.5		
IFWDMN	N	34	15	14		
IFSTMN	M	1.6	1.5	2.2		
IFSTMN	S	0.5	0.5	0.6		
IFSTMN	N	29	18	29		
MATIMN	M	4.6	5.5	7.2		
MATIMN	S	1.6	1.0	0.4		
MATIMN	N	40	29	29		
NBRDMN	M	0.7	1.0	1.0		
NBRDMN	S	0.3	0.0	0.0		
NBRDMN	N	40	29	29		
ATINMN	M	2.1	2.1	2.0		
ATINMN	S	0.9	0.7	0.2		
ATINMN	N	14	28	29		
ATLRMN	M	2.1	2.4	2.3		
ATLRMN	S	0.4	0.4	0.1		
ATLRMN	N	40	29	29		
CLSSMN	M	7.0	6.9	7.0		
CLSSMN	S	0.0	0.1	0.0		
CLSSMN	N	40	29	29		
ROLEMN	M	8.3	8.0	7.8		
ROLEMN	S	0.7	0.5	0.3		
ROLEMN	N	40	29	29		
NSTDMN	M	7.0	7.9	10.8		
NSTDMN	S	2.9	1.9	6.6		
NSTDMN	N	40	29	29		
TIMEMN	M	208.1	329.6	684.0		
TIMEMN	S	81.7	75.9	116.0		
TIMEMN	N	40	29	29		
RNKIMN	M		2.5	2.4		
RNKIMN	S		0.7	0.5		
RNKIMN	N		2	23		
RMKEMN	M	1.8	1.6	2.1		
RMKEMN	S	0.4	0.4	0.6		
RMKEMN	N	38	29	29		
NCISMN	M	1.2	1.4	1.3		
NCISMN	S	0.3	0.3	0.3		
NCISMN	N	40	29	29		

Table 10

Checklists - Spanish:
Descriptive Statistics on the Quantity Indices of Instruction
by Instructional Year for Site 5 Bilingual Sample

Scale	Statistic	IY0	IY1	IY2	IY3	IY4
IFLTPR	M	60.0	80.2	71.1		
IFLTPR	S	26.4	18.5	11.5		
IFLTPR	N	40	29	29		
IFWDPR	M	26.0	6.7	5.1		
IFWDPR	S	20.9	9.2	6.7		
IFWDPR	N	40	29	29		
IFSTPR	M	14.0	13.1	23.8		
IFSTPR	S	12.7	14.0	13.0		
IFSTPR	N	40	29	29		
ATINPR	M	7.0	32.4	33.3		
ATINPR	S	12.3	14.7	8.9		
ATINPR	N	40	29	29		
ATLRPR	M	93.0	67.6	66.7		
ATLRPR	S	12.3	14.7	8.9		
ATLRPR	N	40	29	29		
IFT2PR	M	40.0	19.8	29.0		
IFT2PR	S	26.4	18.5	11.5		
IFT2PR	N	40	29	29		
IFWDRP	M	62.6	30.9	21.0		
IFWDRP	S	23.5	31.9	28.3		
IFWDRP	N	35	20	29		

8)3

Final Report

TEACHING READING TO BILINGUAL CHILDREN STUDY

Volume 7

Language, Literacy, and Instruction:
Integrating the Findings

Wesley A. Hoover, Robert C. Calfee
Betty J. Mace-Matluck

Document BRS-84-R. 1-VII

Preston C. Kronkosky
Executive Director

Southwest Educational Development Laboratory
211 East Seventh Street
Austin, Texas 78701

November 1984

Final Report

TEACHING READING TO BILINGUAL CHILDREN STUDY

Volume 7
Language, Literacy, and Instruction:
Integrating the Findings

Wesley A. Hoover, Robert C. Calfee
Betty J. Mace-Matluck

Document BRS-84-R. 1-VII

Preston C. Kronkosky, Executive Director
Southwest Educational Development Laboratory
Austin, Texas

November 1984

805

There were many individuals and institutions who contributed to this research effort. We wish to express our sincere gratitude to the parents, students, and school personnel who provided the necessary data from which this study is derived.

In addition, several other individuals made valuable contributions to the study, for which we are indebted: Robert C. Calfee, Sylvia C. Peña, and Blanca de Alvarez.

And finally, we wish to thank the local data collectors at the school sites, many of whom remained with the study throughout its duration: Ramiro Barrera, Beatrice Cantú, Irene Cavazos, Carolyn Cruz, María de Obregón, Gloria de Torres, Gigi Galván, Olga Hernández, Irene Méndez, Guadalupe Treviño, Rosalinda Villanpando, and Gloria Villarreal. Their patience, dedication, and hard work helped make this study a reality.

Betty J. Mace-Matluck
Wesley A. Hoover

TABLE OF CONTENTS

Page

LIST OF TABLES

Table | Page

Table | Page

Table | | Page

LIST OF FIGURES

Figure		Page

81.1

PREFACE

In June 1978 the National Institute of Education (NIE) funded the Southwest Educational Development Laboratory (SEDL) to conduct a longitudinal study on the Teaching of Reading to Bilingual Children. Educators and policymakers alike have long recognized that the ability to read is essential for success in school, in work, and in life; yet many children from second-language backgrounds have trouble learning to read in schools today. The majority of these youngsters are from Spanish-language backgrounds and from low income families. Special programs designed to meet the needs of these children are provided in schools, but there is limited research evidence to guide the development, evaluation, and implementation of these programs. This study is intended to provide information that will result in greater insights into what constitutes a favorable learning environment for children from Spanish-language backgrounds, what instructional sequences and events promote successful and efficient learning of literacy skills, and what the language and literacy outcomes of current schooling practices are for a large sample of these youngsters.

The study was conducted during the years of 1978 through 1984. It is a comprehensive longitudinal investigation of the development of reading skills from kindergarten through fourth grade for a representative sample of more than 350 children from bilingual backgrounds, and for smaller samples of children who, on entry into school, were monolingual in English or Spanish. In this "natural variation" study, teaching and learning were carefully documented in field settings at the several sites.

The goals of the study were to (a) describe variations in both English and Spanish language ability of students living in bilingual communities, (b) document prevailing practices in reading instruction for bilingual students, and c) investigate the relations between the instructional program and student achievement for students with differing entry profiles.

Description of the Study

Surveys of the general and school populations reveal an increase in the number of students whose language resources are not an ideal match to the language of the school. An important question for educational practice and policy centers around the school's responsibilities in this situation. Bilingual programs, English-as-a-Second-Language classes, classroom aides, and "sink-or-swim" approaches can all be found in practice today. From limited evidence now available, none of these techniques has emerged as the one best system.

Hispanics make up the largest and fastest growing school-age population today. The demographics for some states show that over the next decade they may constitute as much as a third to a half of the population. In the state of Texas at present approximately one third of the school children are from Hispanic backgrounds (approaching one

million). They are found in virtually ever school district in the state. Many of the school districts in the southern portion of the state serve school populations of which 75% to 99% of the children are from Spanish-speaking backgrounds and, on entry into school, are often limited in their ability to speak English and to profit from instruction in that language. This population is not restricted to the border areas, however. Large urban centers in the state report as much as 20% of their school population from Hispanic backgrounds, with a concentration of some 80% to 90% in certain of their schools.

It is well documented that, in general, children from Spanish-speaking backgrounds, for whatever reason, often encounter difficulty in our nation's schools; they do more poorly on standardized tests than does the general school population, and their dropout rate is high. Bilingual education, in which students are given instruction partially through the home language until they have attained sufficient proficiency in English to benefit from English-medium instruction, has been the principal approach recommended by the Office for Civil Rights to ensure access to equal educational opportunity for these children. Although many individual programs have had considerable success in improving the academic performance of language-minority students, it has not been demonstrated that these programs generally are reducing inequality of educational opportunity on the large scale that was envisioned.

Growth in reading comes about for most youngsters through formal classroom instruction. Understanding the development of reading, and knowledge of the critical variables that determine success or failure, depends on a careful examination of the instructional program -- not just the label over the classroom door, but the program as actually implemented by the classroom teacher.

Educators have raised several issues about the most effective way to help bilingual children become proficient readers of English. These include (a) valid assessment of the student's ability in the languages of the home and of the school, (b) the optimal balance of formal instruction in both languages, (c) the most effective transfer from one language to the other, and (d) bilingual support within the classroom environment. A major thesis of the Teaching Reading to Bilingual Children study is that addressing these issues (and others) requires a comprehensive and ecologically-valid investigation of the linkage between the child's language and the language of instruction.

Design of the Study

To achieve the objectives of the study, considerable attention was given to the selection of schools, teachers and students, to the instruments for assessing language and reading achievement, and to the methods for evaluating the classroom instruction. Each of these topics is discussed briefly below.

Schools, Classes and Teachers

Twenty schools and 200 teachers from six school districts participated in the study. Included are variations in the nature of the reading program (a range from phonics-oriented to meaning-based), classroom organization (some self-contained, others team-taught), and grade structure (the range of grades in the individual school and the extent of cross-grading both vary). The schools differed in size, SES, urbanicity, locale, and makeup of the student body (from medium to high concentration of bilingual students).

Student Cohorts

The study was undertaken in four cohorts or "waves" of students. Three of the cohorts consisted entirely, or in large part, of bilingual students. The first cohort was small (N=40) and of limited generality; the second was somewhat larger (N=80) and covered a slightly broader array of contexts. The third cohort which was both larger (N=200) and broader in its generality, incorporated a number of procedural improvements based on previous experience in the study and included a monolingual English-speaking sample. The fourth cohort consisted of a relatively small sample (N=60) of monolingual Spanish-speaking students.

All of the bilingual sites were from the state of Texas, as were the monolingual English-speaking students. The monolingual Spanish-speaking students were from one site in Northern Mexico.

The original design of the study called for each student to be assessed and observed from entry to kindergarten through exit from third grade. By covering the full range of the primary years, we would be able to examine the transition from "learning to read" through "reading to learn." For students in programs where the initial stages of reading were in Spanish, we also considered it important to determine the transition to competence in English reading.

The original design was in fact implemented for the first two cohorts; some of the students were tracked from first through fourth grade, but most followed the intended design. Due to limited funding in the later stages of the study the last two cohorts could not be followed for the full four years that were originally intended. The bilingual and monolingual English samples from the Texas sites were observed from kindergarten through second grade, and the monolingual Spanish samples from the site in Northern Mexico were observed from first through third grade (the program did not provide a kindergarten).

The monolingual samples were incorporated in the design to aid in validating the instruments for student assessment. Both the English and Spanish cohorts are small and not selected to be fully representative of monolingual populations. Data from these samples will be presented in Volume 3, as part of the discussion on the adequacy of the instruments for measuring growth. The study was designed to study the course of reading in bilingual students, not as a basis for comparing these students with monolingual youngsters. Accordingly, comparisons

between the various samples will not be made in this report, nor do we recommend that others attempt such comparisons.

Language Assessment

Several types of data were collected for each student on English and Spanish proficiency. Each year, early in the Fall and again in the Winter and Spring, teachers rated their students' language skills. Oral language proficiency tests were administered in the Fall of each year. Finally, audiotaped speech samples were obtained monthly on a rotating schedule in three settings: in the classroom, on the playground, and in the home.

Reading Assessment

Several instruments were used to measure reading achievement. Standardized test scores (mostly English) were collected yearly. More detailed information was obtained from a battery of individually-administered "performance based tests" in both English and Spanish. In kindergarten, the Stanford Foundation Skills Test was employed to measure the child's pre-reading skills. From the end of first grade on, the Interactive Reading Assessment System was administered during the Spring of each school year. This instrument provides independent measures of the student's skills in decoding, word meaning, fluency in oral reading, and comprehension. Finally, informal reading inventories were administered throughout the school year.

Classroom Observations and Teacher Interviews

Project staff conducted monthly observations of the reading instruction in each classroom and interviewed the teachers quarterly about their instructional plans. The observation instrument documented staffing patterns, grouping and organization, time allocation, the language of instruction, the character of instruction, the materials and procedures used, and the response of the students. The interviews focused on the teacher's general instructional objectives, as well as the objectives for individual target students. Taken together, these two instruments yield a rich characterization of the classroom environment for the target students.

Student Entry Variables, Classroom Factors, and Reading Achievement

The primary goals of the analyses were to identify the general relationships that characterize variation in these factors and to look for underlying regularities that are associated with success and failure, both in the early stage of reading instruction and in the year-to-year variations.

Documents

This report is one of a series of eight documents contained in the Final Report submitted to the National Institute of Education. A com-

plete list of these documents is provided on the inside of the cover of this report.

The study was a collaborative effort among a number of individuals and institutions. All members of the research team contributed to the thinking, planning, and writing of this series of documents, however, the individual whose name appears first in the list of authors was responsible for preparing the particular document.

Betty J. Mace-Matluck
Wesley A. Hoover
Co-Principal Investigators

Austin, Texas
November 30, 1984

INTRODUCTION

This volume, seventh in a series of eight volumes comprising the Final Report for the Teaching Reading to Bilingual Children Study, explores the linkages between different sources of information contained within the study's multifaceted, complex, and integrated data base. Provided below is a graphic representation of the overall data structure (Figure 1) showing the major elements and the specific sources of information within each element. The sources are listed in order of priority. Those at the top of each list are considered the most valid and informative data sources and are the focus of this report.

Previous volumes have provided a general introduction to the study (Volume 1); described the overall design (Volume 2); discussed the methodology used in the analysis of the data (Volume 3); and presented a discussion of the instruments and the data obtained from them for the bilingual sample, reporting descriptive and summary statistics for each of the data sources within each of the major elements of the data structure (Volumes 4, 5, 6). Reported in Volume 4 are characteristics of the students' oral language entry skills and subsequent growth patterns. Pre-reading skills and reading achievement are discussed in Volume 5. Instruction delivered to the students is characterized in Volume 6.

To assist the reader in assimilating the information contained in the present volume, a brief review of the relevant information contained in the previous volumes is provided below. The remainder of the volume treats the linkage between the various sources of information and is organized around four major sections: <u>Methods of Analysis</u>, <u>Descriptive Results</u>, <u>Integrative Analyses</u>, and <u>Discussion</u>.

Overview of Previous Volumes

The primary goal of the study was to examine the relations between current schooling practices and the language and reading achievement of a large sample of low income Hispanic children who began their initial schooling in bilingual classrooms. More specifically, the study sought to provide information that could result in greater insights into what constitutes a favorable learning environment for children from Spanish-language backgrounds, what instructional sequences and events promote successful and efficient learning of literacy skills, and what the language and literacy outcomes of current schooling practices are for these youngsters.

Growth in reading comes about for most youngsters through formal classroom instruction. Understanding the development of reading, and knowledge of the critical variables that determine success or failure, depends on a careful examination of the instructional program -- not just the label over the classroom door, but the program as actually implemented by the classroom teacher.

PRECURSORS

Language Samples
Stanford Foundation Skills Test
Oral Language Proficiency Tests
Teacher Ratings

Cognitive Style Indices
Cartoon Conservation Scales

INSTRUCTION

Reading and Mathematics Observation System
Teacher Checklist
Attendance

Inventory of Bilingual Instruction
Survey of Teacher Language Skills
Teacher Cognitive Style

ACHIEVEMENT

Interactive Reading Assessment System
Standardized Tests
Informal Reading Inventory

Figure 1. Data structure for the SEDL Bilingual Reading Study.

Educators have raised several issues about the most effective way to help bilingual children become proficient readers of English. These include (a) valid assessment of the student's ability in the languages of the home and of the school, (b) the optimal balance of formal instruction in both languages, (c) the most effective transfer of skills from one language to the other, and (d) bilingual support within the classroom environment. A major thesis of the Teaching Reading to Bilingual Children Study is that addressing these issues (and others) requires a comprehensive and ecologically-valid investigation of the linkage between the child's language and the language of instruction.

To accomplish the goals of the study, more than 300 children in 20 schools in six school districts and taught by more than 200 teachers were tracked from kindergarten through second or third grade (fourth grade in some cases) -- a critical period for the development of literacy.

The study carefully examined the children's language on entry into school and thereafter. Standardized test data were collected and examined, as were other more detailed sources of language and reading data. Systematic observation was carried out in the classroom. Information was gathered about the teachers' instructional plans, and the nature of the instructional program was carefully documented.

The students' entry skills and subsequent achievement in the various components of language and reading were examined, and the students' rates and patterns of growth were investigated in relation to the instruction each student received.

Major Findings: Precursors, Instruction, Achievement

The primary analyses of the data from the study aimed toward four basic outcomes:

- o.. Precursors. Descriptive information using validated precursor profiles typically found in bilingual children on entry to schools throughout Texas.

- o.. Instruction. Class-level descriptions of the approaches used to teach reading to children from bilingual backgrounds in the state.

- o.. Achievement. Development and validation of a set of longitudinal achievement indices that could be used to assess growth in the various components of reading in English and Spanish.

- o.. Linkage. Development and validation of a set of procedures for measuring the linkage between reading achievement on the one hand, and precursor and instructional indices on the other hand, taking into account the possibility of interactions between precursor profiles and response to type of instruction.

For the convenience of the reader, the results of the analyses of the three data sets above (Precursors, Instruction, Achievement), reported in previous volumes, are summarized and discussed below.

Precursors

Language and related cultural differences have been identified as key factors in the educational failure of minority students both by educators and by the courts. Similarly, it is well documented that the student's experiential background and level of social, emotional, and cognitive development at entry into school (i.e., readiness to benefit from formal instruction as provided in the schools) are related to student academic progress. An important question for educational practice and policy centers around the schools' responsibilities in these matters.

To address the issue of language and cultural differences, special language assistance programs, often under mandate by state and/or federal legislation, have been implemented in the schools. The goals of such programs are to concurrently develop English language proficiency while at the same time ensure progress in academic skills achievement. Special assistance, in the form of English-as-a-Second Language classes, has been provided over the years in some schools in an attempt to meet the first of these needs, but it was not until the passage of the Bilingual Education Act in 1968 that schools generally were encouraged to include instruction in the native language of the students to address the second of these (academic progress while acquiring the necessary English skills).

The best means by which to accomplish the above goals have not been clearly established. The nature of the population to be served, as well as local resources and educational philosophies, has given rise to a variety of organizational structures and instructional approaches for the delivery of special language assistance programs. Such variation among programs was observed and documented in the present study. While each of the Texas school sites was subject to the Texas Education Agency's regulations and guidelines, local conditions (e.g., locale; concentration of Hispanic students within the school population; resources; qualifications, experience, and skill of the school's professional staff; perceived needs of the students; extent of use and role of Spanish in the wider community) affected school practices. Notable differences among these were in length of stay of students of particular language categories in bilingual programs, extent of use of Spanish during instruction and in the school environment, point of onset of reading instruction in English for students who began their initial reading instruction in Spanish, and availability of preschool programs.

To address the "readiness for school" issue, Texas, along with several other states, has recently mandated school-based programs for four-year-old youngsters. During the course of the present study, the existence of such programs at the Texas sites was spotty. Those that were in existence were supported for the most part by local funds and were limited in the number of students served.

Oral Language Growth

The students in the bilingual sample were deemed by their schools to be Limited English Proficient, as determined by their performance on a standardized test of oral proficiency given in the Fall of their kindergarten year, and were therefore enrolled in bilingual kindergarten or first grade classrooms when they entered the study.

For the purpose of assessing the students' oral language abilities and monitoring their language growth, three types of language measures were used in the study: (a) an oral language proficiency test, (b) teacher ratings, and (c) audiotaped interactions -- language samples.

The oral language proficiency test used by each of the sampled Texas school districts was the Language Assessment Scales - LAS, except for one year in one school district. The students' entry protocols were collected from the school districts, and both the English and Spanish versions of the LAS was readministered by SEDL staff to the target students in the Fall of each subsequent year. In addition, the teachers were asked to observe the language performance of the target students in their classrooms and then to rate, on the basis of a set of descriptors, the language performance of their students on a five-point scale for each of four language components (pronunciation, grammar, vocabulary, comprehension) and for a fifth category identified as "Overall Communicative Skill." Finally, audiotaped speech samples were taken once a month from selected target students on a rotating schedule in three communication settings: in the classroom, in the home, and either on the playground or in other non-instructional settings within the school.

Analyses of the oral language data suggest the following:

o.. The students in the sample, on entry into school, varied considerably in their degree of bilingualism.

o.. The students, generally, made considerable progress in acquiring skill in English; less growth was observed in their performance in Spanish.

o.. Site differences were apparent in the students' facility in Spanish and in English on entry and in their subsequent growth in each of the languages.

o.. Site differences were also observed in the patterns of language choice, both at entry and over time.

- o.. The students' oral language proficiency varied, in both languages, as a function of the type of measure used (oral language proficiency test versus teacher ratings) as well as by the type of task within a given measure (story retelling versus discrete items that required short-answer responses).

- o.. When compared to teacher ratings, the oral language proficiency test used appeared to underestimate the students' ability in both languages at entry and, at higher grades, to overestimate their English abilities and underestimate their skills in Spanish.

A number of critically important instructional issues surround language assessment. Primary among these is the question of adequate and accurate assessment of the oral language abilities of young children. Objective measures, such as the currently-available standardized oral language proficiency tests, have been widely criticized. The widespread dissatisfaction with these measures arises from the belief that these tests do not reflect the totality of the language resources that children possess, nor do they adequately predict children's ability to perform in the school setting. Further dissatisfaction arises from the concern that formal testing of young children's language may in fact be measuring many things other than language (e.g., general readiness for school; knowledge of test taking). Subjective measures, such as teachers' ratings, have been maligned by some who point to the "human element" that comes into play with such procedures. Natural, or free speech, samples avoid some of the potential pitfalls of other types of measures, but they, too, have their limitations.

Issue: Valid language assessment. The SEDL research staff, fully aware of the limitations of the various kinds of measures and of the hazards involved in oral language assessment (given the state of present knowledge about what constitutes oral language proficiency and how to assess it), employed multiple measures in an attempt to obtain a reasonably accurate index of each student's oral language abilities and patterns of language choice over time. Analyses of the oral language data strongly suggest that none of the existing measures by themselves provide adequate information on which to base educational decisions. Use of a variety of measures and procedures can, we believe, provide a reasonably accurate index of the student's oral language abilities. However, this process is time consuming and requires skill and expertise that often is not readily available within most school districts. Given that results of oral language assessment figures prominently in a number of educational decisions regarding schooling practices for language minority children (e.g., identification, program placement, termination of special services), further research is urgently needed to determine not only effective but practical means for assessing the oral language proficiency of young students.

Issue: Language and program placement. When examining language as a precursor skill for reading achievement, additional instructional issues emerge. First, to what extent does the child's language at the time of entry into school determine program placement? By legislative mandate, all children in Texas from non-English language backgrounds who, at entry into school, score at or below a predetermined cutoff score in English on the district-selected oral language proficiency test (the LAS in most cases) are placed in a bilingual education pro am, which implies some use of the home language (e.g., Spanish) for instruction for some given period of time. Thus, the issue here is not one of access to the program but rather the accuracy and adequacy of the information on which placement decisions are made.

Issue: Language and actual program. Second, to what extent does the child's language at the time of entry into school determine the actual instructional program delivered? The teacher's perceptions of the child's language abilities and instructional needs determine to a large extent the instructional treatment delivered to the student. Therefore, even in bilingual classrooms, use of the home language for instruction for a given child or group of children will vary, both as a medium of instruction and for support within the classroom environment. Length of stay in the program is also determined to a large extent by the teacher's perception of a student's readiness to perform in an all-English classroom, as well as by prior instructional treatment and the student's progress in acquiring English. In the present study, the majority of the students (approximately two-thirds) did not receive any reading instruction in Spanish, although all were enrolled in a bilingual program when initially selected for participation in the study. With the exception of one site which showed rapid exit from bilingual reading programs, students who did receive Spanish reading instruction were likely to stay in those reading programs for at least two years.

Issue: Language and reading achievement. Finally, to what extent does the child's language at the time of entry into school affect subsequent reading achievement? Two issues are involved here: choice of language for initial reading instruction; relationship between oral language development and reading acquisition.

It is generally believed that reading is a single process and that having learned to read in one language, reading in another known language is a matter of transferring and extending one's existing knowledge and skills. It is also generally believed that bilingual children learn to read more easily and more efficiently when their initial reading instruction is provided in their stronger language. Therefore, transitional bilingual education programs may provide initial reading instruction in Spanish for children who are clearly Spanish dominant and are limited in their English skills at the point that formal reading instruction is begun (usually in first grade). In the present study such instruction was provided for approximately one third of the students. While all of the students in the sample were deemed by their schools to be Limited English Proficient on entry into school and were enrolled in bilingual classrooms when they entered the study (as kindergarteners for the most part), subsequent placement and instructional decisions resulted in initial reading instruction in

English for the majority of the students. Some students were transferred to a regular mainstream program at the end of their kindergarten year, presumably because they had either made rapid progress in acquiring English and were therefore no longer considered Limited English Proficient or because their English skills had been underestimated in their entry language assessment. Other students in the sample scored low in both languages (or were perceived by their teachers to have attained less than adequate oral language development in either language), and therefore it was presumed that English reading instruction would be as appropriate for these children as would Spanish reading instruction. Yet other students who remained in bilingual classrooms in first grade and received initial reading instruction in English may have gained sufficient skills in English to begin such instruction but were deemed in need of support in the home language in other curriculum areas. Thus, contrary to popular belief, not all children enrolled in bilingual classrooms receive reading instruction in their non-English home language.

The literature is replete with studies that have shown a moderate-to-strong relationship between oral language development and reading achievement. Knowledge of the language being read is at the heart of the reading process. Reading is a derived skill in that it builds upon oral language and requires the translation from writing to a form of language from which the reader already is able to derive meaning. To learn to read, children must bring their knowledge of the spoken language to bear upon the written language. A well-developed system of oral language assumes a functional vocabulary and the ability to discover the structure and meaning underlying spoken utterances. It also assumes a rudimentary ability to reflect upon language that allows children to discover the properties of spoken language that are central to the correspondence between its written and spoken forms (e.g., awareness of segments of speech at the levels of word and subunits of words, awareness of relationships among words in text, as well as among higher-order structural units such as clauses and sentences). Children who do not have a well developed understanding of the communicative process at entry into school often experience difficulties in learning to read and therefore fall behind the school's expectations in academic progress. Well designed preschool programs, along with parent involvement components of school programs, have gained support as a means of enhancing the language development of young children.

Pre-reading Skills

Underlying the general skills that are critical to acquiring new knowledge and skills (e.g., attention, memory, verbal fluency, effects of previous learning) are a set of independent component skills that are intimately related to the acquisition of reading. These include decoding, word meaning or vocabulary, sentence and paragraph comprehension, and text comprehension. Assessment of the students' pre-reading skills at entry into the present study revealed the following:

- The overall sample of students came to school with sufficient skills to begin literacy acquisition -- they did not appear to be academically disadvantaged.

- Approximately one-half of the sample of students came to school knowing the letter names of the English alphabet, which has been found to be a good predictor of early English literacy exposure.

- Knowledge of the Spanish alphabet was negligible, but expected, given its different treatment in the language and culture.

- Sight-word recognition was minimal in both languages, but higher in English than in Spanish.

- Visual matching skills were already highly developed.

- Auditory segmentation skills could readily be acquired with familiar words by most of the students, with higher performance with English words than with Spanish words. The transfer of this skill to novel items was difficult for some.

- Vocabulary knowledge was high, with slightly greater strengths in English.

- The formal dimensions of schooling and text (as measured by listening comprehension) appeared to be new to many.

- Transferability across languages is suggested by the following correlational patterns in the data: (a) visual matching tasks and the metalinguistic task of auditory-phonetic segmentation possess a degree of transferability between the two languages. (b) the linguistic tasks tapping vocabulary knowledge and comprehension are independent <u>across</u> (but not within) the two languages, and (c) alphabet knowledge and sight-word recognition tend to be related across the two languages.

<u>Issue: Pre-reading skills and reading achievement</u>. To what extent does the child's pre-reading skills development at entry affect subsequent reading achievement? The correlational patterns between student performance on the pre-reading measures and the reading achievement indices can be summarized as follows. First, knowledge of the English alphabet at kindergarten entry was found not only to be generally related to English literacy skills at first-grade entry, but also to subsequent growth in decoding and reading acquisition. Knowledge of the Spanish alphabet, however, did not carry such widespread predictive power for Spanish literacy development, neither for entry skill nor for subsequent growth. Kindergarten skill in decoding and oral language was related to such skills at first-grade entry, within both English and Spanish; but for English, some of these entry skills

were further related to subsequent English literacy growth (segmentation to decoding growth, and oral language to reading growth).

Clearly, children's knowledge about literacy at entry into school has an impact on their reading achievement both in the early stages of literacy acquisition and in later reading achievement. An important question for educators is, "Can instruction change the relative level of attainment in literacy that is predicted by individual differences between children in their knowledge about literacy on entry into school?" A number of studies have shown that differential progress in the acquisition of literacy is related to the quality of the instruction delivered to children. The problem remains, however. Children who are well prepared at entry to take advantage of what the school has to offer progress at the rate of at least a year of growth for a year of instruction; children less prepared often get off to a slow start, and even if they progress at the rate of a year of growth for a year of instruction, they still lag behind their more advantaged peers as they progress in school. A challenge for the schools is to find means for helping the less academically advantaged children become better prepared to benefit from instruction and for accelerating their growth in the early years so that they can keep pace with the general school population of their same age.

Instruction

A coordinated system of classroom observations and teacher interviews provided rich and extensive data on the instructional program each target child received over the course of the study. The major findings concerning the dimensions of instruction assessed in the study are summarized below.

Of the 250 bilingual students in the sample, 70 students (28%) began reading instruction in Spanish. Of those, some received Spanish reading instruction for one year before being transferred to English reading; others remained in Spanish reading for two, three, or four years, with most being transferred to English reading by the end of third grade.

Analysis of the instructional data suggests that the teachers generally implemented the instruction they had planned, as indicated by a close match between the instruction observed and what teachers said that they were going to do during a given period. Instruction in Spanish and English was similar in terms of the instructional dimensions assessed in the study.

In general, the instruction delivered may be characterized as follows:

o.. The size of the instructional groups for reading ranged from about 13 in the early years to about 15 in the later grades.

o.. The teacher (as opposed to an aide or some other "instructor") was associated with the target students about two-thirds of the time.

o.. The role played by the teacher tended to be one of facilitation (rather than direct instruction) in the aggregate.

o.. The language of instruction tended to be English during the English reading classes and Spanish during the Spanish reading classes, but with some English used during the Spanish reading period.

o.. About half of the instruction time devoted to reading instruction during the first and second grades was focused on decoding; this fell to about 30% by fourth grade.

This instruction tended to involve non-explicit letter-sound pairings for each of the grades (e.g., children were shown a word on a flash card and were asked, "What does this word say?" or were asked, "What is the name of this letter?"); little explicit letter-sound work was observed (i.e., direct instruction in or practice on tasks that required the child to overtly focus on a specific isolated linguistic element and relate it to its graphic representation).

o.. The amount of time spent on developing word meaning was small.

o.. About 30% of the instructional time was on instruction in the meaning of sentence and texts (comprehension) in the first three years, with a slight increase in the fourth year.

The quality of this instruction was fairly stable across years, and was generally non-explicit (e.g., favoring a focus on literal facts over making inferences).

o.. Independent work accounted for about half of the instructional time during the first two years, dropping to about 35% in the following years.

o.. The level of formal language demand (i.e., the extent to which the activity or task required interaction with connected instructional text either oral or written) observed in both independent and group work was low, starting at a relatively low level in the first year, and increasing to mid level by the last two years.

o.. The primary materials used in instruction tended to be basal readers accompanied by workbooks, worksheets, and chalkboard/paper/pencil activities.

- The number of nonengaged students was low; productivity was rated medium each year, while noise tended to be low.

Typically what was seen when observing in the classrooms, viewing the observation protocols, and reviewing the teachers' instructional plans was the traditional basal reading lesson conducted in small groups and strongly influenced by the content, sequencing, and pacing of the teacher's manual.

The summary data on which the above descriptions were based were subjected to factor analysis in order to reduce the number of instructional indices. In both the English and Spanish data sets, seven factors were derived. The seven factors identified in the analysis of the English observational data included the following:

- Engaged Text Time, an index of reading time where students were engaged with text materials.
- Direct Group Instruction, an index of direct instruction delivered by an instructor and which was aimed at groups of students, rather than individuals.
- Quality of Formal Language, a measure of the formal language demands made upon the students.
- Amount of Decoding, a measure of the relative amount of time devoted to instruction in decoding.
- Productivity, an index of the conditions promoting high individual student productivity.
- Secondary Materials, a measure of relative usage of secondary materials (though this interpretation is oversimplified).
- Number of Students, an index of the number of students constituting an instructional group.

The factor solution derived from the Spanish data was similar to that derived from the English data, with five of the seven factors containing many of the same component variables. The seven factors identified in the analysis of the Spanish observational data included:

- Quality of Formal Language (corresponding to the third English factor derived).
- Direct Group Instruction (the second English factor).
- Engaged Text Time (the first English factor).
- Number of Students (the last English factor).
- Amount of Decoding (the fourth English factor).

o.. Secondary Materials, an index of both the quality and quantity of secondary material usage (only tangentially related to the sixth English factor).

o.. Control, a complex factor without an English correspondence, which is essentially an index of the number of management interruptions.

Educational research over the last 10 to 15 years, conducted primarily with students from the general school population, has produced a well-founded knowledge base that allows educators to point with confidence to characteristics and actions that differentiate between instructional settings in which students successfully master the learning goals set out for them and those in which students are less successful. It identifies and describes what effective teachers do and how effective instruction is accomplished in effective schools. Similarly, some of the most eminent reading experts claim that the best teachers in the best schools know how to turn students into proficient readers.

Research in bilingual education and related topics has also accumulated a substantial knowledge base in the last 10 years. The focus on academic achievement prior to and in the early part of the decade led to more and more investigations into the interaction between differences in the languages of instruction and the language of the student. This in turn uncovered a variety of variables which led to research into school and classroom climate, teacher and student variables, and pedagogical, socio-cultural, and legal issues. Thus, a considerable body of research exists that speaks directly to issues related to language minority education.

When examining the findings from the instructional data from the present study in relation to the knowledge base contained in the literature referred to above, one finds instructional patterns and teacher behaviors that are associated with (a) student academic gains in monolingual classrooms, (b) successful practices in bilingual classrooms, and (c) less reading gains in monolingual classrooms.

Factors present in the data that are associated with student academic gains and successful practices in both monolingual and bilingual classrooms include the following:

o.. Strong focus on academic work; time spent working with textual materials (as opposed to time spent with puzzles, games and toys).

o.. Time allocated to reading and academic verbal interaction; 20 to 30 minutes allocated to reading group direct instruction.

o.. Use of active teaching practices; a great deal of instruction from and interaction with the teacher; time spent in small groups (as opposed to one-on-one); interactive on-task instruction; student engaged time.

o.. High achievement expectations; use of tasks of appropriate difficulty level that challenge the students but which allow consistent success.

o.. Efficient classroom management; allocated instructional time devoted to instruction; classrooms that are relatively free of major behavioral disorders.

Additional factors associated with successful bilingual classrooms include the following:

o.. Use of the home language with Limited English Proficient students some of the time.

o.. Use of English primarily during English-medium instructional periods and Spanish primarily during Spanish-medium instructional periods.

Factors associated with less student gains in reading in the present study and in research on monolingual students may be summarized as follows.

o.. Amount and quality of decoding instruction (inappropriate amounts or timing of such instruction; non-explicit instruction on letter-sound pairing); lack of evidence of teaching children strategies for achieving pronunciations that are both systematic and correct, helping children understand the rules of and significance of syllabification for decoding, assisting children in their efforts to blend sounds to produce syllables or words.

o.. Limited attention given to explicit instruction to develop vocabulary (word meaning) and higher-order comprehension strategies (beyond those of comprehension of literal facts).

o.. Ability grouping of students, which may not be in the best interest of low achieving students. Children who get assigned to the lower groups may get locked into an instructional track in which the range of instruction delivered is such that these students have limited opportunity to learn more than a narrow range of the skills and content needed to become fluent readers.

o.. Extensive use of seatwork assignments for low reading group students. Recent research suggests that seatwork is qualitatively a different experience for lower achievers than for high achievers. The two groups differ in terms of fluency of their answers and the appropriateness of strategies used. This may explain why achievement difference widens over time. Low achievers are spending less of their seatwork time in beneficial ways.

Issue: Quantity and quality of instruction. To what extent does the quantity and quality of the instruction delivered to bilingual children affect reading achievement. Of the many factors that impact student progress in reading, instruction is the one factor for which the schools have primary responsibility and over which they have the most control. Therefore, identifying instructional patterns that are associated with success and failure, both in the early stage of reading instruction and in subsequent years, is a critical issue surrounding improvement of practices for all children.

Issue: Nominal instructional program. To what extent does amount of reading instruction in Spanish (i.e., the number of years of reading instruction in Spanish prior to entry into exclusive English reading instruction) delivered to Limited English Proficient Hispanic students predict reading achievement in English? The issue of choice of language for initial reading instruction was discussed above (Issue: Language and reading achievement). The focus of the issue here is the amount of instruction (or the duration of the instruction) for children to receive maximum benefit from such instruction. In the Texas schools in the sample, criteria for transition from bilingual programs to all-English medium classrooms included a specified level of achievement in Spanish reading. In general, the children who remained in the Spanish reading programs the longest were children who were having difficulty in learning to read. Thus, one would expect an inverse relationship between number of years of reading instruction in Spanish and reading achievement.

Student Attendance

Patterns of attendance in school have been shown to be related to achievement. Children from agricultural migrant families are often withdrawn from their home schools in the month of April to accompany their families as they move from place to place to harvest seasonal crops. Typically, these children have little or no additional formal schooling until they return to their home schools, often a month or so after school starts in the Fall. Approximately one-third of the target students in one of the sites in the study was from migrant families.

Attendance at the each of the sites was relatively stable. Most absences were short-term and were usually related to illness or family emergencies.

Issue: Attendance and achievement. To what extent does attendance affect reading achievement. Except in the case of the migrant children, attendance at school was relatively normal and, therefore, would not be expected to be a strong predictor of reading achievement in the present study.

Reading Achievement

A primary purpose of the study was the investigation of patterns of growth in reading achievement. The study employed multiple measures for assessing each of the major components of skilled reading (vocabulary knowledge, decoding, and text comprehension), and for the bilingual sample, monitored such growth in both English and Spanish. Presented below are summaries of the data obtained from two of these sources: the Interactive Reading Assessment System (IRAS) and standardized achievement test scores.

Interactive Reading Assessment System (IRAS)

The IRAS, an individually administered diagnostic assessment system designed for research application, provided the primary index of reading growth for each student. Modelled after the informal reading inventory, the IRAS provides independent measures of several component skills essential for fluent reading.

The materials in the test were selected to cover a wide range of skills and knowledge in the areas of reading and oral language from the level usually expected of a mid-year first-grader to that of a junior high school student. The areas of knowledge assessed in the system include: reading of isolated words, definition of common words within and beyond the student's reading vocabulary, and selected word analysis skills based on the pronunciation of synthetic words. Comprehension of connected text is also assessed in several contexts: reading and listening comprehension of both narrative texts (typical of those found in reading texts and literature series), and more difficult expository texts.

The materials within each subtest are ordered by difficulty based upon grade-level expectations of performance, with each IRAS level roughly corresponding to a half-grade level. Thus, materials contained within the fourth level of a given subtest correspond to materials which average second grade students should be able to read.

Testing with the IRAS was done in the Spring, and all target students were tested beginning in first grade and continuing until exit from the study. All target students in the bilingual sample were tested with both the English and Spanish versions.

Since entry language skill was found to be significantly related to subsequent student growth in the various components of reading, as measured by the IRAS, the data are summarized below both for the overall sample and for subsamples of students broken out by language skill at entry, as determined by teacher ratings. Performance in English is treated first followed by a summary of student performance in Spanish. Relations within and between the IRAS measures is then summarized as are those between the pre-reading measures and the IRAS measures.

IRAS-English.

o.. For the overall sample, the students entered first grade with English oral language skills which exceeded the expectations of the growth track model, but which grew in accord with the model predictions; thus oral language skills were above grade-level expectations throughout the primary grades.

o.. The decoding skills of the students were minimal at first-grade entry, and showed subsequent growth which was above grade-level expectations (progress in spelling, however, was slow); thus decoding (of isolated words) was, like oral language skill, above expectations throughout the primary grades.

o.. Decoding fluency may have presented problems in reading connected text, as by second-grade exit, the average student had a reading rate of less than two syllables per second.

o.. Reading comprehension was about half a grade level below expectation at entry, and showed growth slightly above the expected rate; thus reading comprehension was found to be slightly below grade level expectations throughout the primary grades.

o.. Students who came to school with relatively lower English skills showed greater growth in English oral language capacities, and showed a convergence in such skill in late fourth grade with those students who entered with higher English skills. However, the high English entry students were better able to profit from decoding instruction in that their initial advantage in decoding continued to expand.

o.. Students with relatively higher Spanish oral language skills at entry into kindergarten had growth rates in English reading comprehension which exceeded those of students with relatively lower entry Spanish oral skills. This suggests that although the development of English listening comprehension did not differ for these groups, relatively higher skills in Spanish at school entry promote the growth of English reading comprehension.

o.. A significant relation exists between entry level English skills and reading performance in English. The low English group began first grade just below grade level expectation, with subsequent growth that gave about three-quarters of a grade-level improvement for each year of instruction. The high English group began first grade slightly above grade-level expectations, and grew at a rate that was slightly below grade-level expectation. Thus, by fourth-grade exit, the high English entry group was projected to be about a half grade level behind, while the low English entry group was projected to be more than a full grade level lower.

IRAS - Spanish.

- o.. In the overall sample, the students entered first grade with Spanish oral language skills which exceeded the expectations of the growth track model, but which grew at half the expected rate; thus oral language skills were above grade-level expectation at entry, but were projected to fall below grade-level expectations during the primary grades.

- o.. The decoding skills of the students were minimal at first-grade entry, and showed subsequent growth which was slightly above grade-level expectations (as in English), thus decoding (of isolated words) was above expectations throughout the primary grades -- progress in spelling, however, was slow.

- o.. As in English, the data suggest that decoding fluency may have presented some difficulty in reading connected text.

- o.. Reading comprehension was a grade level below expectation at entry, and showed growth which was only half the expected rate; thus, reading comprehension was found to be substantially below grade-level expectation throughout the primary grades.

- o.. Entry level skills in Spanish were related to reading performance in Spanish. The low Spanish entry students entered first grade with less skill than the high Spanish group in the areas of formal language and decoding, but subsequent growth did not differ. For reading comprehension, however, the two groups began with the same low-level skills, but, given the greater formal language and decoding skills of the high Spanish entry group, their growth in reading comprehension was able to proceed at a greater rate. This rate, however, was substantially below that expected from the growth track model, and the data suggest that the major difficulty for these students was not decoding skill, but rather, skill in dealing with the formal language aspects of text.

Relations within and between IRAS measures (English and Spanish).

- o.. Within both English and Spanish, the relationships found between the component scales can be summarized as follows. The highest relationships are generally between the component scales within the three major skill areas assessed (formal language, decoding, and reading). The correlations between these skills are strongest for decoding and reading, somewhat weaker between formal language and reading, and weakest betwe - decoding and formal language. Thus, the general correlational pattern suggests that decoding and formal language skills are relatively independent, with both needed for growth in reading comprehension.

o.. The correlational pattern between the English and Spanish indices is as follows: skill in decoding, and to a lesser degree, in reading, are related across the two languages, while formal language skills (as expected) are generally unrelated. There is, however, a general trend for stronger relationships between a given English task across the set of Spanish tasks when compared to those relationships for the same given Spanish task across the set of English tasks -- this suggests that literacy development in English may be more readily transferable to Spanish than from Spanish to English.

Relationships between pre-reading and reading achievement measures.

The correlational patterns between the pre-reading measure and the reading measure (IRAS) is as follows:

o.. Knowledge of the letter names of the English alphabet at kindergarten was found not only to be generally related to English literacy skill at first-grade entry, but also to subsequent growth in decoding and reading acquisition. Knowledge of the letter names in Spanish, however, did not carry such widespread predictive power for Spanish literacy development, neither for entry skill nor for subsequent growth.

o.. Kindergarten entry skill in decoding and oral language was related to such skill at first-grade entry, within both English and Spanish; but for English, some of these entry skills are further related to subsequent English literacy growth (segmentation to decoding, and oral language to reading growth).

o.. For the cross-language correlations, few significant relationships were found.

Standardized Reading Achievement Tests

Generally beginning in first grade, standardized achievement tests were administered to all students in the Texas sites in the Spring of each year. Three different standardized tests were used over the course of the data collection phase: the California Achievement Test, the Comprehensive Test of Basic Skills, and the Iowa Test of Basic Skills. Standardized achievement tests in Spanish were not administered systematically, nor to any great extent, by any of the schools in the study. Performance of the students on the standardized reading achievement tests in English may be summarized as follows:

- o.. Performance in English indicates that the students in the overall sample entered first grade just slightly below grade-level expectations and showed growth which was also slightly below expectation -- by fourth grade exit, the sample was projected to be a full grade level behind.

- o.. High English entry students were projected to be about a half grade level below expectations at fourth grade exit, with the Low English entry group projected to be reading at about the third grade level -- two grade levels below expectation.

Issue: Site characteristics and reading achievement. To what extent do socio-cultural factors within the local community affect reading achievement? In addition to issues related to language and instruction discussed above, factors associated with community environment undoubtedly played a role in shaping the growth patterns of the students. These include, but are not limited to, locale (proximity to the U.S.-Mexico border and the community's orientation to communities across the border), geographic isolation, and concentration of Hispanics in the school and community. They also include the extent to which the two languages are used in the community, as well as the role of the home language in the affairs of the home and of the community; attitude of the student and others toward the maintenance of Spanish and/or the acquisition of each of the languages; and the extent to which written materials and formal usage are available to the students in each of the languages.

The remainder of this volume addresses the various issues identified above. More specifically, it assesses the degree to which various entry skills and instructional program indices can account for above or below average skill in each instructional year with respect to the set of component reading skills which were of primary interest (decoding, listening comprehension, and reading comprehension).

METHODS OF ANALYSIS

Overview

The purpose of this section is to describe the methodology employed to integrate the several data sources that were obtained in the study. The section begins with a consideration of the nature of the analytic problem. On the face of it, "integrating the findings" has a simple intuitive meaning, but this goal actually can be interpreted in several ways. Next we consider various methodological strategies that might be taken in approaching the task of integration, and the strengths and difficulties of each.

The major portion of the section comprises an account of the method that has actually been employed, together with a discussion of the advantages and disadvantages of this particular methodology. Briefly, for what is best viewed as a preliminary analysis of an extraordinarily complex and challenging data base, we have used a "rough and ready" method that allows us to gain some sense of the structure of the data. The primary outcome measures are all from the Interactive Reading Assessment System (IRAS). This instrument yields several measures in both English and Spanish. For each year that a student was tested, a deviation was computed between each of the student's IRAS measures and the aggregate growth track index summarizing average performance during that year. The IRAS deviations for each year were then submitted to a regression analysis in which the predictors were oral language classification on entry to school, previous year's performance on the corresponding IRAS measure, nominal program category or the number of years of Spanish reading instruction, instructional dimensions as measured by RAMOS and the Checklist, attendance, and site at which the school was located.

For statistical convenience, most of these indices were standardized by year. Factor scores were used to summarize the instructional variables, and linear contrasts to represent the categorical variables. The primary goals of the analysis were (a) to determine the degree to which the several predictor indices were consistently related to the outcome variables and (b) to evaluate the structural patterns of any such relations.

The advantage of this approach is its methodological simplicity and the attendant reasonableness in time and computing costs. The disadvantages are a loss of information about the time course of student growth and problems that arise from multicollinearity among the predictors. Since one of the primary design features of the study was the emphasis on investigating the nature of student progress over time, the fact that the present analysis does not take advantage of this feature is regrettable. The decision to terminate data collection for the primary cohort of students when they were in second grade seriously undercut the longitudinal value of the study, and so the choice of analytic method is only partly consequential in this regard. The problem of multicollinearity in this data base is not as serious as it might be, and a more careful examination and refinement of the data structure might reduce the extent of multicollinearity to negligible proportions. The resources available for analysis of the data structure were quite limited, however, and primary attention has gone into assuring the trustworthiness of the data base (substantial cross-checking was done) and adequate documentation of the information for purposes of secondary analysis.

Despite the roughcut character of the method employed in this volume, the findings have proven quite informative, as will be seen in the sections that follow. Variations in student achievement are predictable from the information gathered in the study, and the patterns of performance are related both to the precursor factors (including measures from the previous year), but also to variations in the instructional program. The present effort at integration is best viewed as yielding insight into the informativeness of the data base; more sophisticated analyses, building on this foundation, can certainly provide a more precise and detailed account of the specific effects of instruction on performance.

The Analytic Question

What does it mean to "integrate the findings" from the study? The previous volumes have described the hypotheses that motivated the investigation, the design for identifying the sample of districts, schools, classrooms, and students, and methods of instrumentation for documenting various dimensions of growth in language and reading achievement along with the precursor and instructional factors that were identified as candidate variables for explaining student growth. The descriptive findings for each constellation of measures have been presented; we view many of these as of substantive importance in their own right.

In its original conception, the study held the vision of a "natural experiment." By locating the appropriate array of districts and school sites, it would be possible to assess the effects of different approaches for handling the needs of students who were bilingual on entry to kindergarten or first grade. If this vision had been realized, then integrating the findings might have been relatively straightforward, using the standard experimental model as a paradigm. The precursor measures, in particular oral language classification, would provide an index of aptitude. The instructional measures would confirm the nature of the different treatments. Any of several analytic techniques could test the two primary questions that arise from this design:

> Do the different treatments lead to differences in achievement that are statistically trustworthy and practically substantial in the aggregate?
>
> Do the effects of the treatments vary as a function of student aptitudes, such that there are interactions between treatments and aptitude that are statistically trustworthy and practically substantial?

The research literature provides a rich array of examples of this genre, using either multiple regression methods (e.g., Cronbach & Snow, 1977) or analysis of variance with block assignment (e.g., Calfee & Piontkowski, 1984). If this approach had proved workable, "integration" would be most simply represented by the two questions posed above.

As it turned out, the "natural" world of the schools did not in this instance provide a clearcut foundation for experimental comparisons. As was seen in Volume 2, the bilingual programs in the sample were not clear exemplars of well defined hypotheses of any sort. Specification of the program by districts and schools was at a fairly general level. Implementation of these general policies varied from school to school, from teacher to teacher, and from year to year. The "treatments," in short, proved to be highly variable.

In addition, instructional and situational factors not directly related to the type of bilingual program clearly emerged as significant candidates for explaining variation in student growth. Some of these factors are clearcut: the amount of time allocated for reading, the curriculum objectives, the teacher's organization of the class for instruction, the management of the class, the type of materials, and so on. Other factors were more difficult to isolate because of situational confoundings. Students are not assigned to either schools or classes at random. As a consequence, students at some sites were more likely to be in class environments where (a) many of their peers were more or less fluent in Spanish or English, (b) teachers were or were not bilingual, and/or (c) the program did or did not place an emphasis on parallel instruction of some sort in both English and Spanish. In short, the need was for a method of describing the students' instructional environment as a multi-dimensional array of partly confounded factors.

Then there is the question of how best to assess student outcomes. The standard answer to this question has typically been through reliance on standardized achievement tests. We have foregone this route for several reasons, including our conviction that these instruments are not directly representative of reading skills and knowledge, but are also influenced by overall adjustment to school and test-taking skills, among others. IRAS was chosen as an alternative more directly reflective of the various components that make up literacy. While we think that this choice was justified by the results, two complications are attendant on it. First is the fact that the several components, because they are not highly correlated and hence do not reduce to a single factor, call for some type of multivariate analysis.

Second is the question of how to deal with changes in performance over time. IRAS measures were designed to follow a linear growth track, and the design was generally effective in realizing this design goal in the aggregate. In an ideal world, the growth track model could have been used to derive two indices for each of the IRAS components for each student: estimated level at time of entry to school, and rate of growth over the elementary school years.

Unfortunately, the simplest form of the growth track model can be rejected as an adequate representation of student performance for two reasons, one of which is artifactual and can be remedied, the other of which is substantive and might prove valuable as a basis for a more sophisticated assessment of the outcomes of the study. As to the first matter, all of the measures were bounded at both the top and bottom of the scale, the latter of necessity (if a student cannot read at all, there is no way of extending the scale below "zero"), the former because of limited resources (more difficult materials could have been created for the more able students in the later grades, but we simply did not have the staff and time for this task). These bounds do not show up in the aggregate charts in Volume 5, but they do appear in the data for individual students with sufficient frequency to compromise the growth track model. These boundary effects can be dealt with in several ways. The upper-limit effect can be handled by extrapolation (once a student has reached the top of the scale, assuming that progress is relatively linear up to that point, then project performance beyond the ceiling). The lower-limit or floor effect can be reconciled in a similar fashion, but with less confidence about the meaning of the procedure (presumably one is assessing varying degrees of "reading readiness" or the lack thereof by projecting below the zero point on the scale for a given IRAS index).

The second reason for nonlinearities over years in the IRAS indices is more substantive in character. The growth track concept can be viewed as a particular way of representing a learning curve. As such it is based on a number of assumptions. One of these is that the set of items that comprise each of the IRAS components comprise materials of constantly increasing difficulty; it is the constancy that leads to a linear change in performance. The method of choosing the materials to obtain constancy was strictly empirical, based largely on the vocabulary counts that are also employed by the publishers of basal reading series in the design of instructional materials. Thus, the differences between successive levels of IRAS follow the publishers' guidelines. This strategy has a practical basis, and entails no assumptions about "what might be." The fact that most of the IRAS components change linearly in the aggregate over grades speaks to the success of the strategy in achieving the design goal.

Another assumption of the growth track concept is the constancy of the instruction provided to the individual student. That is, suppose that a student progresses through the various levels of a basal series at the normative rate suggested by the publisher. Then we would expect progress on the various IRAS components at a steady straightline rate. We know, however, that not every student progresses at the normative rate. A detailed accounting of differences in student progress through the basal curriculum was documented as part of the study; another result of the limited funding provided for data analysis is the decision to forego this type of analysis for these data, and the analysis of the informal reading inventory protocols that were obtained to match the curriculum progress.

In any event, data from the study show that whereas some students made normative progress, and a few moved at a faster rate (a rare occurence because of the implications for the teacher in the next grade), many students did not move through the materials laid out in the basal series at the expectation for a given grade. Other research shows that completion of a given level of material, quite apart from the student's performance on entry to a given level, is one of the strongest predictors of achievement at the end of a period of training (Barr & Dreeben, 1983). The consequence of variations in progress through the curriculum is the presence of perturbations in the linearity of the growth track. A student may not show any gain in decoding performance during a particular grade because he or she did not spend any significant amount of time studying letter-sound correspondences, followed by a year with a teacher who stressed phonics and consequent gain in achievement -- to the degree that such patterns exist in the data, then the simplest application of the growth-track model in formulating the analytic question is undercut.

Given all these considerations, the question of how to integrate the various data sources is by no means a simple one. Viewed in the broadest perspective, we have a multivariate and longitudinal data structure for a relatively large sample of students with differing profiles on entry to school, who experienced a multivariate and longitudinal set of experiences during the three to five years during which our investigators were able to obtain data about these youngsters. To repeat, not all of these data are being considered in these analyses. Nonetheless, the most fundamental question can be posed as follows: What relational patterns, if any, describe the linkages between experiences and outcomes over time? The methodological challenge is interesting. On the one hand, a simple aptitude-treatment approach runs into the hurdle of dealing with the multidimensional and longitudinal character of treatment, aptitude, and outcomes. On the other hand, an ethnographic approach (dealing with the data from each student as a particular configuration) foregoes the strengths of the study as a generalizable design in terms of the sample and the instrumentation; it is not clear how one can aggregate the descriptive impressions that derive from an ethnographic examination of the patterns.

The study staff have considered several ways to resolve the conundrum sketched above. We are of the opinion that analytic methods exist for realizing the full potential of the data, and some alternative approaches will be considered in the next section. All of these are demanding of substantial time and other resources, and so had to be set aside in favor of a more pragmatic approach. Nonetheless, it has been interesting to consider "what might have been."

Alternative Methods of Analysis

The tendency, when "blueskying" a problem, is to roam rather broadly. We will restrict this section to two topics. The first deals with methods for dealing with the outcome data for each student as a configuration rather than a collection of data points. The second speaks to the task of reducing the multicollinearity of the set of predictor factors so that they also can be dealt with as a set of relatively distinctive configurations.

Configurations of Reading Performance

As noted above, average performance in reading achievement is not necessarily typical of any individual. This generalization is of basic importance in all data analysis, to be sure. The researcher is obliged to report not just averages but also standard deviations; the people inhabiting a day care center may be twenty years old on the average, but two-thirds of the people are under four while the others are over thirty, with no one close to twenty!

The IRAS scores present a similar challenge in description, but with an additional dimension because of the longitudinal character of the data. This matter was discussed in some detail in an interim report (Calfee, Mace-Matluck, & Hoover, 1982, pp. 79ff); the material in that report should be consulted for a more complete discussion of the issue. Reading scores vary from the averages reported in Volume 5 in several ways. The averages trace out a growth track that increases with grade in school, and that is predominantly linear. Individual students may differ from the average in that (a) their intercept is higher or lower than the average, (b) their rate of growth is higher or lower than average, (c) both of the above are observed, or (d) their growth track is significantly nonlinear, in which case the estimates of intercept and slope are not well defined.

In the 1982 Interim Report, profiles for several students were presented to demonstrate these variations. The exhibits spanned a range of possibilities. As mentioned above, boundary limits are responsible for some artifacts in the data. A few students, by the end of first grade, were at the upper limit of one or more IRAS components. The instrument does not allow measurement of growth for such students. Figure 2, reproduced here from the 1982 Interim Report demonstrates the point. This student, by the end of first grade, was extraordinarily fluent in Spanish, and performed at the upper limit of the test in defining and decoding single words.

A second noninformative pattern of growth can be best described as idiosyncratic. Student 0007, for instance, had an inexplicable problem in decoding (English) synthetic words when tested at the end of second grade. The drop of almost three grade-equivalents from the overall trend makes no sense in terms of expected growth trends, and is probably not a valid index of the student's actual level of competence. To be sure, IRAS was designed to provide measures of achievement based on actual performance, but it nonetheless entails a limited sample of data. Ideally, assessment includes multiple sources of information about a given domain of skill and knowledge; the judgment that this student's performance on decoding synthetic words is not being validly assessed at the end of second grade is supported by the parallel assessment of decoding of regular vocabulary, which shows a steady pattern of growth. Multiple assessments, which were provided in the design of IRAS, are valuable for evaluating idiosyncracies, but they also provide confirmation for the internal consistency of an assessment system.

The profile for student 0007 also illustrates what might be labeled as plausible nonlinearities in student progress. Notice that this youngster's growth from first to second grade in most areas approximates the normative expectation. English Definitions, Vocabulary Decoding, Reading and Listening Comprehension, and Spanish Comprehension, on average, increase from the end of first grade to the

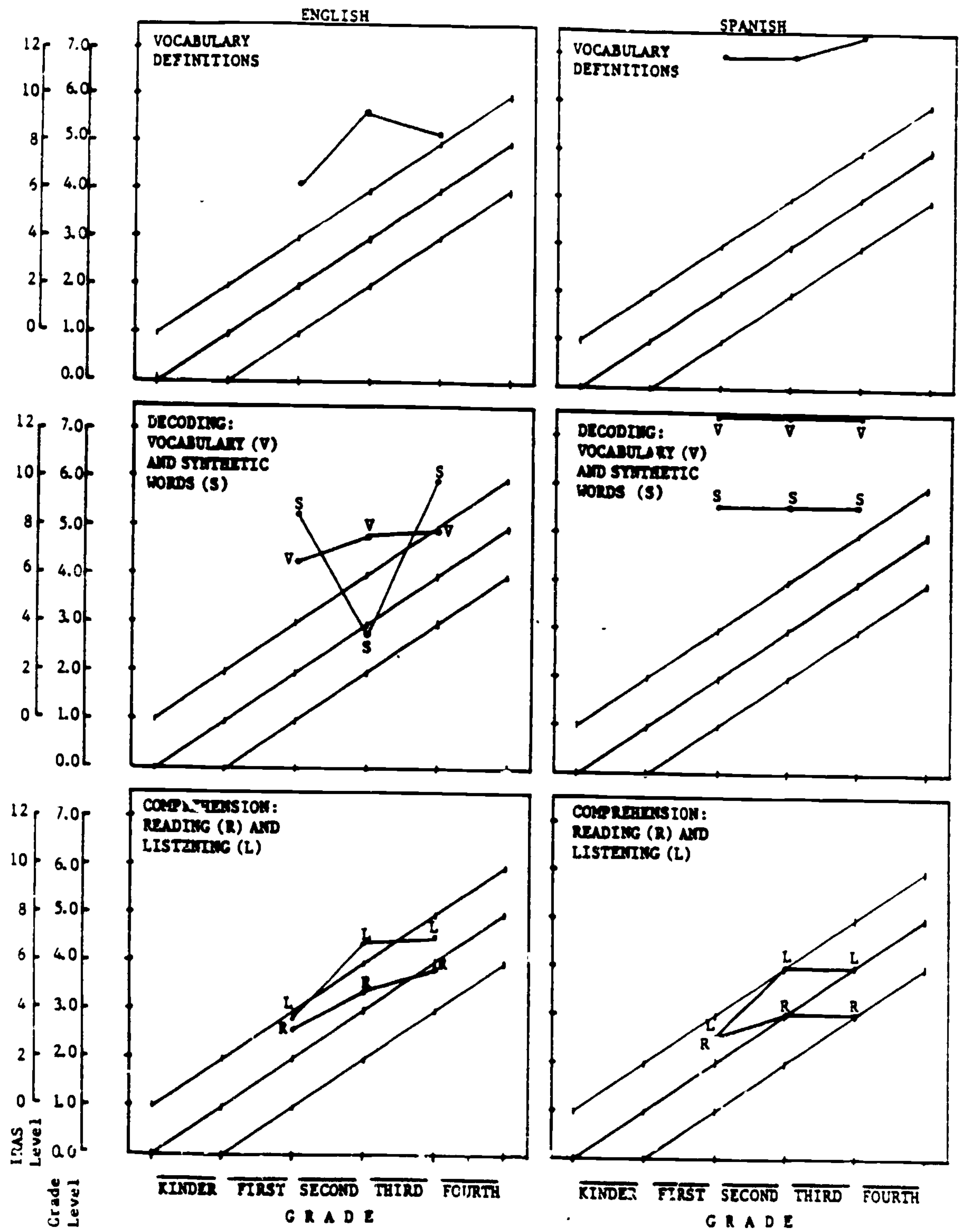


Figure 2. IRAS performance profiles in English and Spanish for Student 0007 (Group A).

end of second grade by about one grade-equivalent on the IRAS scale. These same indices, however, increase by little or nothing from the end of second grade to the end of third grade. The lack of growth in these measures is not reflective of a ceiling effect; there is room for improvement on all of these measures.

A related but distinctive pattern appears in the protocol for student 2097 (Figure 3), also from the 1982 Interim Report. For virtually all of the English IRAS measures, and for the Spanish IRAS measures that emphasize decoding, this student shows no evidence of growth during first grade, nor from the end of first grade to the end of second grade. During third grade, however, there is a substantial improvement in performance, consistent over virtually all the IRAS indices and of remarkable extent, so that by the end of third grade the student is performing more or less at grade level.

We need at this point to reiterate a limitation in the data set. All of the protocols that were examined in the Interim Report were from students who had been assessed three times -- at the end of first, second, and third grades. Most of these students had also been assessed at the beginning or end of kindergarten, but (necessarily) on different measures. With three data points, it is meaningful to speak about idiosyncracies and other nonlinearities. As noted above, reduced funding for the study meant that for the third and primary cohort, students were assessed only at the end of first and second grade. Accordingly, there is a limit to the value of this data set for evaluating configurational patterns. As Rogosa, Brandt, and Zimowski (1982) have pointed out in their work on change, so-called two-wave data cannot answer many fundamental questions about the nature of change; more to the point, they cannot answer many fundamental questions about the relation between treatment factors and extent of change.

Against the background provided by these examples, and in light of the character of the data set, we can now present the concept of a configurational analysis, and indicate why we have chosen to pursue an alternative for the present volume. First, let us note that if the assumptions of unlimited and constant growth for individual students had been consistent with the observed data, then analysis of the entry point and growth rate parameters of the growth track model would have exhausted the infromation available about individual differences in student progress. But examination of the student protocols reveals that in many instances there appears to be consistent departure from the simple "constant progress" model. We are limited to those students for whom assessment was available for three or four grades, but even so, the number of consistent "inconsistencies" of a meaningful sort is sufficient to suggest that the model does not tell the entire story.

One type of systematic variation is exemplified by the protocol for Student 0007 shown in Figure 2, where initial growth is followed by a plateau. More typical is the protocol of Student 2097 in Figure 3, where little or no movement from the baseline appears until the end

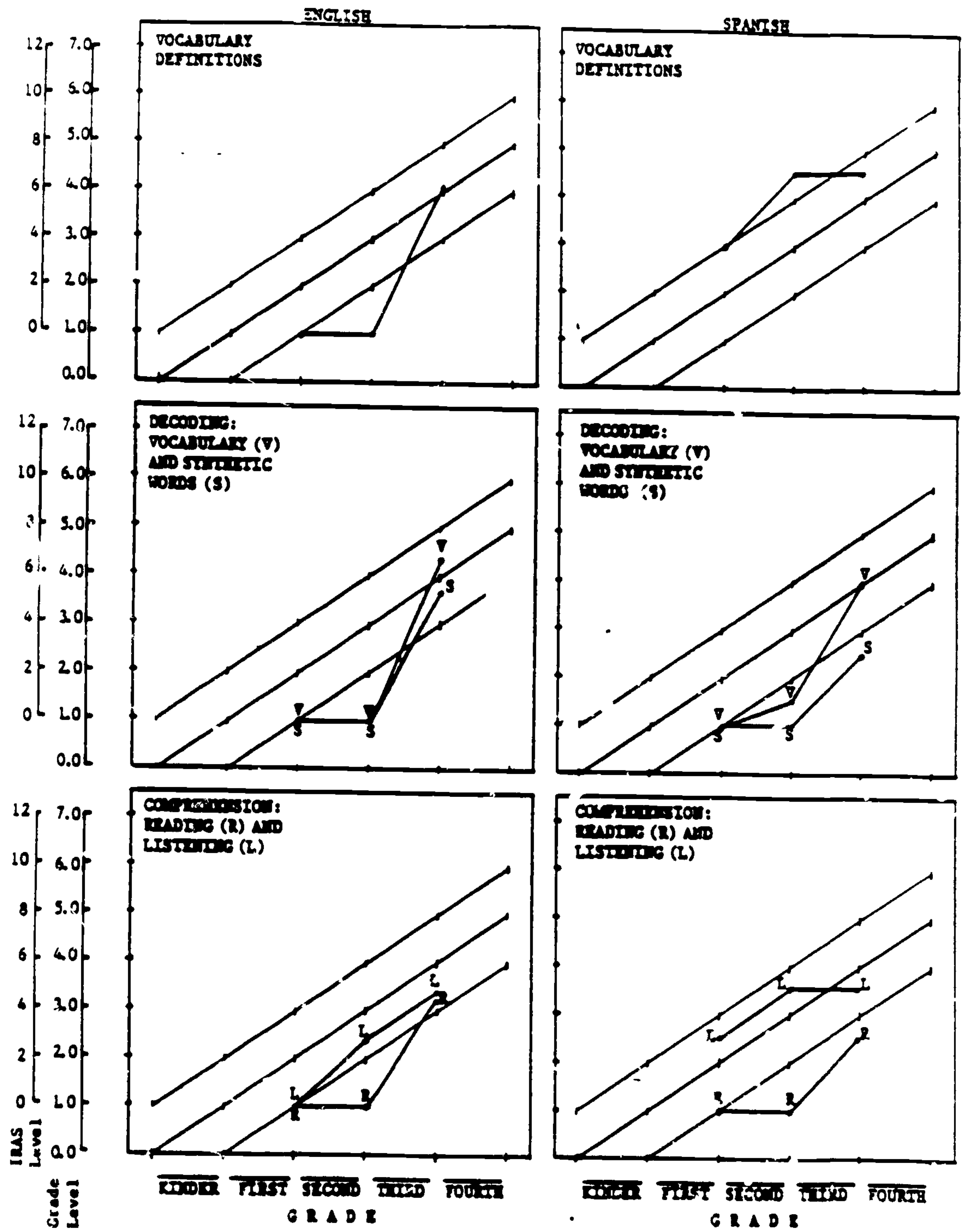


Figure 3. IRAS performance profiles in English and Spanish for Student 2097 (Group C).

of second or third grade, followed by a spurt in performance. These patterns are sometimes apparent in only one or two of the IRAS components; at other times several of the components follow the same trend.

One can venture various hypotheses about the cause for these trends or patterns. A "readiness" hypothesis might hold that the changes reflect developmental effects, especially for those students for whom progress is delayed for a year or two past entry to first grade. An "instructional" hypothesis would look to the events in the classroom as a cause. If the student is provided decoding instruction neither in English nor in Spanish, then it is understandable that the student might make little progress in those areas of reading that require this skill. This situation might arise because of the nature of the curriculum materials provided to the teacher, because of the character of the other students in a class, or for a variety of other reasons. One can even combine the two hypotheses; the student's developmental level might lead the teacher to delay systematic instruction in the abstractions of decoding.

On the other hand, decoding instruction might be provided, but in a situation that was instructionally ineffective. The teacher might have trouble managing time or the classroom environment; the program might rely on worksheets and seatwork; the materials might be at too easy or too difficult a level to promote learning; and so on.

Other hypotheses might be entertained. The point of configurational analysis is that departures from average performance trends (in the present instance, the linear growth track that holds on the average for most IRAS components) call for an explanation. The most immediate source of the explanation lies in the instructional program, for which departures from the average are also available. Thus, the positive or negative departure of a student's performance on an IRAS index can be compared with the departures of various RAMOS and Checklist indices from the average. In the 1982 Interim Report, we compared individual student protocols with corresponding instructional protocols, and found several instances of interesting matches: for students who showed little growth in reading in first and second grades but made substantial progress in decoding English during third grade, the third grade program of instruction, compared with first and second grade, appeared in these instances to rely more on direct instruction, to stress English, and to employ texts and other printed materials in a systematic fashion.

Where the data are sufficient, one can compare any two patterns by calculating a least squares measure of the match. In this instance, the patterns are changes in relative performance over time compared with changes in relative program emphasis over time. Each student's IRAS protocol represents a pattern that can be represented as fitting the standard (linear growth) or showing some degree of discrepancy. To the degree that one or more of the instructional program indicators shows a match to the discrepancy (i.e., a close least-squares match), then we have an indication that the program indicator is a factor determining changes in student progress. This method would, in principle, provide a sensitive technique for understanding the influence of specific program components on student achievement. The examples in the 1982 Interim Report were provocative in this regard.

We have decided against pursuing this methodology for several reasons. One set of problems has to do with the development of a technology for carrying out the procedure. Although the approach is conceptually straightforward, it is not immediately obvious how to carry out the pragmatics. We are not aware of any standard statistical routines for computing the measures, nor is there a standard for assessing statistical trustworthiness of the method.

A more serious problem arises because of the limitations in the data. The largest cohort of students was assessed at only two points in time; the configuration approach is only workable when three or more data points are available. To be sure, ancillary data were collected that could in principle provide a richer pattern of change over time, but this information was not included in the analysis, again because of limited resources. Our inability to pursue the configurational approach is regrettable. The typical approach of evaluating student achievement at one point in time by precursor and treatment factors, which will be pursued here in a later section, is subject to a number of limitations to be noted. The configurational method represents, in theory at least, an innovative approach to a question that remains unexplored to the best of our knowledge: What is the effect on learning of changes in the program during learning? The so-called "N=1" designs found in a number of studies of behavioral modification treatments verge on this question, but in these studies both the treatments and the outcomes are typically unidimensional so that the question of "match" reduces to a matter of change in performance upward or downward.

Orthogonal Program Configurations

In order to determine the contribution of a given factor on an outcome measure, it is important to separate this influence from other competing factors. To the extent that two factors are correlated or collinear, the investigator will experience difficulty in assessing the separable effects of either factor, much less the interaction of the two. To the degree that the researcher can exert some influence in the assignment of subjects to various treatment conditions, it is possible to arrange a design in which collinearity or confounding is kept to a minimum. In "natural" experiments, this degree of control is impossible. Moreover, natural processes are often such that confoundings are commonplace, and such is the case in this study. The extent and character of confoundings in our data will be discussed later; for now, suffice it to say that data are subject to a substantial degree of multiple collinearity (i.e., sets of predictor factors that are mutually related to one another).

When the conditions of a study do not permit control over confoundings by random or systematic assignment, the researcher can sometimes employ analytic methods that alleviate the collinearities. Analysis of covariance is one such method, a. is post hoc blocking. These methods are not entirely failsafe, and uncritical reliance on them can lead to faulty interpretations. Nonetheless, when used with caution they can aid in separating the influence of correlated factors.

In the present study, the array of factors is sufficiently complex that the simpler methods are not immediately applicable. An alternative, which was used with considerable success in a previous investigation of Chinese bilingual students in Seattle (Hoover, 1983), is to refashion the set of factors into a collection of nested orthogonal linear contrasts. This method, like any other, works only when the degree of confounding is relatively modest.

The method can be illustrated by considering two factors that are typically confounded in practice: verbal ability and the child's gender. Boys in the primary grades tend to perform more poorly and girls more highly on tests of verbal aptitude. Suppose the goal is to determine the effects of aptitude and gender on reading achievement. The confounding of the two predictor factors can pose a barrier to the clear identification of the influence of the two factors. One can, however, normalize verbal aptitude within boys and girls, so that the aptitude measure now is indexed relative to gender. The redefinition of aptitude does not solve all problems of interpretation, but it can present a clearer analysis as a basis for interpretation.

The same principle can be applied to any set of correlated factors. The basic approach is to carefully examine the structure of the set of predictors, determine the relative priority of the several variables within a confounded set for assessment of the research hypotheses, and then create a set of nested contrasts that are mutually independent. The process is tedious, and requires an intimate familiarity with the entire data structure. The payoff is that the analysis leads to more clearcut results, both for evaluating the main effects of high-priority factors, and for assessing potential interactions among factors.

In the absence of such a reconfiguration of a set of correlated predictors, the researcher perforce relies on the computer algorithm to distribute predictable variance among the several variables. Generally speaking, little provision exists for determining the presence of interactive effects. In essence, the standard methods for conducting a multiple regression analysis in the presence of collinear predictors take the form of a battle of sorts; to the degree that any group of predictors contains substantial intercorrelations, the measured relation of any particular predictor to the outcome measure is likely to reflect peculiarities in the distribution of errors in both the predictor and outcome variables. Deletion or addition of a small percentage of cases may yield a substantial alteration in the pattern of results.

In our opinion, the use of nested orthogonal contrasts for analysis of the present data set would entail substantial advantages in the clarity of the findings. Unfortunately, the amount of time required for this analysis and the costs in additional computer runs made this effort impossible.

Present Method of Analysis

This section describes the methods for analysis of the relation among the several sets of variables listed previously: precursors, instruction, and achievement. Briefly, each IRAS component in English and Spanish was converted into a set of year-by-year deviations from the aggregate growth track for each student. The result was a set of 72 outcome indices: nine IRAS components, four years, and two languages. Each of these was submitted to a multiple regression analysis in which the predictors included entry language classification, previous year's performance, nominal reading program, instructional dimensions, an attendance index, and site contrasts.

In the remainder of this section, we will describe the methods for computing each of the measures, the sequence of steps used in the analysis, and the advantages and limitations of the approach. The descriptions will be sketchy; they are designed to give an overall picture of the methodology, which will be presented in greater detail in the sections that follow.

Reading Achievement Outcomes

The Interactive Reading Assessment System seemed to us to provide the most complete and workable set of measures of reading achievement. Standardized test scores were available for virtually all students (in English), and analyses of these measures were reported in Volume 5. The scores are of limited dimensionality, however, and the tests varied over sites. At the other extreme, informal reading inventories were obtained on a regular schedule during each year of the study. These data, which provide a rich source of information about student progress, entailed many challenges in analysis, and a complete examination was not possible, given the available resources.

Each administration of IRAS yields the nine component measures listed below:

ORAL LANGUAGE	WORD/SENTENCE DECODING	READING COMPREHENSION
. Vocabulary Definition	. Vocabulary Decoding	. Narrative Reading
. Narrative Listening	. Letter-sound Decoding	. Expository Reading
. Expository Listening	. Letter-sound Spelling	
	. Sentence Reading	

All but two of these components were designed according to the growth-track concept, according to which the expected value at entry to first grade is a score of zero, after which progress occurs at a rate of two IRAS levels per grade. For Letter-sound Spelling, the measure was the percentage of words spelled according to conventional letter-sound correspondences; for Sentence Reading, the measure was the number of syllables per second.

In the present analysis, performance for each student was represented by eight vectors of the nine measures listed above: deviation scores for each of the four years, one set for Spanish and the other for English. The deviation was computed simply as the difference between the IRAS scale score and the corresponding point on the aggregate growth track computed across all students with at least two valid scores on a particular combination of component and language.

Precursors

Two types of precursors were incorporated in the present analysis: oral language level on entry to kindergarten and previous year's achievement in reading. Oral language level was determined for both English and Spanish as a two-level category based on teacher ratings. As will be remembered from Volume 4, teachers rated student language in the Fall and Spring of each year, including kindergarten. These ratings coincided quite well with the oral language samples; measures on the LAS correlated with teacher ratings and with the language samples, but appeared less trustworthy than the teacher ratings. We would have preferred to use the language samples as a precursor index of language competence, but these measures generally were available for only a subsample of the students.

For both languages, the students were divided at the median rating category from the Fall kindergarten rating. For English the median rating was 3.0, while for Spanish the value was 4.0. As it happened in this sample, oral language competence was virtually independent on entry to kindergarten, as evidenced by the chi-square between the two median splits, X^2 (1, $\underline{N}$ = 254) = 3.05, $\underline{p}$ < .10. Analyses of variance were conducted for each of the IRAS linear longitudinal measures with English and Spanish language rating as the independent factors. The results showed that English IRAS measures were generally affected by the English language rating, and Spanish IRAS measures were generally affected by the Spanish language rating. The effects of the complementary language rating were generally negligible, as was the interaction between the two language ratings. Accordingly, in all of the regression analyses, oral language as a precursor was represented by the corresponding language rating around the median split.

A student's achievement level at the end of a given school year is generally related to performance at the beginning of the year. Accordingly, an index of previous performance was included as a precursor. From second grade on, the corresponding IRAS deviation served as the index. For first grade achievement, the Alphabet Knowledge subtest from the Stanford Foundation Skills Test was employed. For English, awareness of the letter names is known to be correlated with later reading achievement, for reasons that are not entirely clear. The distribution of scores on the subtest is bimodal in this sample, replicating earlier findings, and so the precursor was scored as a dichotomous contrast.

Program and Instruction Factors

Two measures of the student's status in a bilingual program were selected for the present analysis. As indicated in earlier volumes, program status was often difficult to determine, and for many students assignment to bilingual programs was limited to the first year or two of schooling. The first program index was the total number years of assignment to Spanish reading instruction of any sort. This index, an estimate based on several data sources in addition to RAMOS and Checklist notes, was coded as follows:

0 No Spanish Reading Instruction
1 Kindergarten Spanish Reading Instruction Only
2 K-1 Spanish Reading Instruction
3 K-2 Spanish Reading Instruction
4 K-3 Spanish Reading Instruction
5 K-4 Spanish Reading Instruction

The index reflects the number of years in which there was some evidence of assignment to a Spanish reading program; it does not indicate the intensity of the program nor any other characteristics.

The second program index was a dichotomous variable indicating whether the student was assigned to Spanish reading instruction during the particular instructional year. The two program variables are necessarily correlated.

Instructional factors included the summary scores derived from the factor analyses described in Volume 6. There were seven RAMOS summaries and five Checklist summaries for each instructional year. English instructional summaries were used in the regression analyses for English IRAS deviations and Spanish summaries for Spanish IRAS deviations. Additional information on these measures is provided in a subsequent section of this volume which presents descriptive statistics.

Miscellaneous Predictors

Attendance data were provided every year by the districts for each student. The absentee rates were quite low in general, but given the emphasis on "time on task," we decided to include this variable in the analysis. The index was the percentage of days attended during the instructional year.

The study was conducted at six different districts or "sites," which were selected to represent a wide range of demographic variations. Some of the differences between sites are reflected in other predictor factors (e.g., language on entry to school, Spanish reading program indices, instructional features). In order to assess any other between-site effects that were not included as part of other influences, three orthogonal contrasts were introduced as the last step in the regression analyses:

SITE05 The contrast between the border and nonborder districts of the state.

SITE35 The contrast between the less urban and the more urban of the two nonborder districts.

SITE12 The contrast between the border districts based on degree of urban influence.

Steps in Analysis

Once the decision was reached as to the general strategy for handling the overall data structure and the variables to include in the analysis, the data were examined in a series of well defined stages. The first stage entailed a review of the descriptive features of the various indices. Most of the variables had been studied in some detail for the preparation of the reports in previous volumes in this series. The present analysis required transformation of some of the variables, however, and so the additional checks were warranted. Where possible, steps were taken to restrain the deletion of cases because of "missing data." For instance, in the RAMOS and Checklist summaries, certain categories were relatively rare and hence might appear as missing. The average value was substituted in these cases, and a "missing factor" index was incremented to provide a check on the effect of the substitution.

The next stage was the examination of the correlational patterns among the variables. We first considered correlations between the IRAS deviations. If all of the outcomes were highly correlated, then there would be little point in conducting a large number of redundant analyses. Next the degree of collinearity among the predictor variables was examined. While the findings point to the desirability of a more sophisticated approach, the situation was reasonably workable from our perspective.

The last stage preparatory to regression analysis was the examination of correlations between the predictors and the outcome measures. Two measures were computed: the zero-order correlation, and the partials with language classification and previous year's performance as covariates. The descriptive data for all of these variables and their relationships are presented in a subsequent section (Descriptive Results).

855

Finally the series of multiple regressions was conducted. IRAS deviations were the dependent measures in each analysis. Predictors were introduced in the order of presentation listed below:

Precursors	Language Classification Previous Year's Performance
Instruction	Years of Spanish Reading RAMOS Summaries Checklist Summaries Attendance Site Contrasts

The regression findings are quite extensive. Our emphasis for this volume is on the proportion of variance accounted for by the various sets of predictor factors, rather than on the presence of "statistically significant" effects; given the large number of analyses, the latter are relatively less interesting.

Variations over Grades

In viewing the results presented in subsequent sections, the reader needs to remain aware of certain features in the design of the study. The original intention was a full-fledged longitudinal design. Because data collection was curtailed for the final cohort, the longitudinal plan is truncated. Some variations in the instrumentation also deserve note. These cautions will be repeated in later sections; the presence of redundancy is worthwhile, given the risks attendant on overlooking these variations.

As may be recalled from previous volumes, the sample comprises three cohorts of students. The first cohort included kindergartners and first graders who were followed through fourth grade. The second cohort consisted of a sample of kindergartners and first graders who were followed through third and fourth grades, respectively. The third cohort, the largest and most representative sample, began the study as kindergartners; the data collection phase of the study was terminated when they exited second grade. The first and second cohorts are from border communities; the first cohort served for pilot work on the instruments during the early stages of the study, when it was viewed as essential to try out the instrumentation with youngsters who were clearly bilingual. The third cohort was drawn from non-border districts. The practical implication of this design is that the third and fourth grade data are from much smaller samples, and from border sites only.

Several changes in instrumentation also occured during the first year or two. Most of these entailed refinements designed to improve the reliability and workability of the methods for monitoring instruction and assessing student outcomes, these refinements were generally implemented in a fashion that provided continuity of the measures. Three of the IRAS components, however, were not part of the battery during the first three years of the study: Sentence Reading, Expository Listening Comprehension, and Expository Reading Comprehension. These changes have relatively modest effects on the regression analyses, but merit note nonetheless.

Finally, the analyses to be presented in this volume are limited to those students in the sample that are classified as bilingual. We have reason to believe that all of the students in this category in fact had some degree of oral language fluency in both English and Spanish. An additional sample of students from a monolingual English background was assessed during the study to determine the appropriateness of the instrumentation for such students. One further sample of students from a site in Mexico permitted evaluation of the instrumentation in this context. Both of these samples were small in number, and of limited representativeness; accordingly, neither merited inclusion in the present analysis.

Advantages and Limitations

As noted at the beginning of the section, the strategy chosen for an integrative analysis of the various data elements is admittedly a "rough and ready" approach, reflecting the resources available for examination of the data and the need for a preliminary assessment of the degree of patterning in the data. As for any strategic decision, this approach entails both advantages and limitations.

The chief advantage is that the approach is feasible within the resources available for analysis. The method of addressing one measure at a time is conventional, the techniques are readily available using standard programs, and the intepretation of the findings poses little in the way of new challenges to most researchers. The method also sidesteps several problems that would beset other more sophisticated strategies. For instance, the longitudinal data string varies with the cohort; more "grades" are available for the earlier than the later cohorts. Variations in the availability of data on the full range of IRAS measures and predictors from previous years are handled as "missing data" for a single analysis.

The disadvantages are at the same time obvious and ephemeral. Because each outcome measure is handled in isolation from the others, there is a loss of information about the structure of reading achievement for the individual student. Because each year is considered in isolation from the others, there is a loss of information about the longitudinal character of changes in reading achievement. Likewise, there is a loss of information about the configurational patterns relating changes in instruction to changes in achievement.

While it is easy to describe the more obvious discrepancies between the original goals of the study and the outcomes that can be realized from the present analyses, it is more difficult to determine the impact of these discrepancies. As will be seen in later sections, the findings from this preliminary work appear to be of value in their own right. There is still debate about the effects of schooling apart from the contribution of the home and other nonschool influences;

previous achievement generally accounts for most of the predictable variance in present achievement. The findings from this study further controvert this generalization, and extend our knowledge about the effects of schooling. But a number of questions will remain unanswered from the analyses, questions that might be illuminated, in our opinion, by further examinations of the data.

DESCRIPTIVE RESULTS

This section presents the descriptive statistics for the variables employed in the integrative analyses, providing summaries of the IRAS deviates and the predictor variables. For descriptive data presented in earlier volumes, the results will only be briefly summarized here; descriptive results for variables presented for the first time will be more thoroughly discussed. As in earlier volumes, English and Spanish reviews will be given separately, treating English first. The discussion begins with the IRAS deviates, giving their associated descriptive statistics and inter-correlations.

IRAS Deviates

As discussed earlier in this volume, deviate scores were calculated for each target student for each instructional year for each of the nine IRAS scales within each language version of the test. These were derived by subtracting the individual target student's actual score from that based on the aggregate growth track (see Volume 3 for a discussion of the growth track model). The descriptive statistics for the deviates are discussed below.

Descriptive Statistics

Table 1 presents the descriptive statistics for the yearly deviates obtained for each scale of the English IRAS for the bilingual sample. The left margin defines the nine scales, and for each, the mean (M), standard deviation (S), and number of cases (N) are provided. The scale names are mnemonic, and stand for the following:

VDC: Real Word Decoding (Vocabulary Decoding)
VDF: Vocabulary Definitions
LDC: Synthetic-word Decoding (Letter-sound Decoding)
LSP: Synthetic-word Spelling (Letter-sound Spelling)
SRD: Sentence Reading
NRC: Narrative Reading Comprehension
ERC: Expository Reading Comprehension
NLC: Narrative Listening Comprehension
ELC: Expository Listening Comprehension

Since the aggregate growth track for a given scale is defined by the average slope and intercept values obtained over the best-fit line computed for each target student through the available data points, deviate values will not necessarily sum to zero (as they would if

Table 1

Interactive Reading Assessment System - English: Descriptive Statistics on Deviates by Instructional Year for the Bilingual Sample

Scale	Statistic	Instructional Year 1	2	3	4
VDC	M	0.04	-0.10	-0.58	-1.69
	S	2.41	3.43	4.37	4.07
	N	249	248	93	58
VDF	M	0.04	-0.04	-0.68	-0.38
	S	2.71	2.49	3.18	3.22
	N	248	246	92	58
LDC	M	0.05	0.14	-0.58	-2.09
	S	2.09	2.42	2.31	2.14
	N	246	246	93	58
LSP	M	0.29	0.50	-3.83	-8.83
	S	13.96	20.24	23.17	22.78
	N	247	242	91	58
SRD	M	-0.07	-0.02	-0.42	-0.80
	S	0.65	0.97	1.02	1.11
	N	184	229	93	58
NRC	M	0.36	0.13	-0.60	-1.32
	S	1.40	2.25	2.65	2.32
	N	249	247	93	58
ERC	M	0.32	0.07	-0.65	-1.12
	S	1.28	2.33	2.83	2.60
	N	182	229	93	58
NLC	M	-0.04	0.06	-0.82	-1.61
	S	2.12	1.96	1.98	1.52
	N	248	247	93	58
ELC	M	0.14	-0.13	-0.85	-1.69
	S	2.17	2.29	2.13	1.79
	N	185	229	93	58

these constituted deviates from the mean of the actual sample of data points within a given scale, language, and instructional year). And indeed, zero values do not obtain, as the general aggregate trend in Table 1 is for small positive deviates in the initial years and larger negative deviates in the later years.

This pattern partially reflects the treatment of floor and ceiling effects. As discussed in Volume 5 (pp. 73-76), successive initial floor or final ceiling data points for individual target students' longitudinal profiles were removed in the computation of the individual linear growth indices. The effect of these data point deletions with respect to positively sloped growth functions is to increase the estimate of the slope, and concomitantly, lower the estimate of the Y-intercept. Since these adjusted growth indices were employed in the computation of the aggregate slope and intercept measures, individual target students' deviations computed from each actual data point (without deletion of the previously removed floor or ceiling data points) will give overestimates in the early grades and underestimates in the later grades, just as seen in Table 1.

A second contributing factor to this trend derives from the cohort structure of the study. As described in Volume 3, differences in aggregate growth indices were associated with the number of constituent data points. Since 60% of the linear estimates were based on data from cohorts tracked only through second grade (and thus providing only two data points per target student), and since the slope estimates were generally higher for such two data point cases as compared to those based on three or four data points, the larger negative average deviates found in Instructional Years 3 and 4 would be expected.

Table 2 presents the descriptive statistics for the deviates obtained for each scale of the Spanish IRAS for the bilingual sample. The format of this table matches that of Table 1 which presented the English summaries.

First, note that the pattern found in the English data does not hold in the Spanish data. Instead, the aggregate picture is one where deviates are of zero or small positive value in the early grades, increasing positively in magnitude in the later grades, especially in the scales concerned with the reading of connected text. Of course these data are subject to both the floor-ceiling effects and the cohort structure effects described above. However, an additional factor is influential here. As discussed in Volume 5, the slopes of the growth functions in Spanish reading were greatly reduced from those found in English, primarily due to a larger proportion of target students who had very limited success in Spanish reading at each instructional year (due to their limited exposure to Spanish reading instruction). As a result, the slope of the aggregate growth function is based on a larger proportion of zero slope components, and thus, the influence in the aggregate deviate indices of students successful in Spanish reading is increased, especially in the later grades where

Table 2

Interactive Reading Assessment System - Spanish:
Descriptive Statistics on Deviates by Instructional Year
for the Bilingual Sample

Scale	Statistic	Instructional Year 1	2	3	4
VDC	M	0.05	0.07	0.73	-0.31
	S	3.40	5.07	5.50	5.36
	N	249	248	95	58
VDF	M	-0.12	0.07	0.81	0.29
	S	3.47	3.86	3.37	3.37
	N	232	247	95	57
LDC	M	0.12	0.1[illegible]	0.46	-0.23
	S	1.63	2.02	1.89	1.77
	N	244	249	94	58
LSP	M	-0.30	0.65	4.44	-3.61
	S	18.76	24.45	24.22	24.04
	N	250	244	93	56
SRD	M	-0.02	-0.08	0.31	0.36
	S	0.49	0.89	1.16	1.20
	N	186	230	95	58
NRC	M	0.60	0.20	0.68	1.27
	S	0.64	1.41	2.28	2.77
	N	250	249	95	58
ERC	M	0.64	0.20	0.52	1.16
	S	0.37	1.27	2.21	2.92
	N	186	230	95	58
NLC	M	-0.04	0.05	1.23	1.05
	S	1.95	2.24	1.66	1.45
	N	248	247	95	58
ELC	M	0.00	-0.04	1.30	1.74
	S	1.75	2.25	1.96	1.65
	N	186	230	95	58

861

relatively more non-zero values are found for those students enrolled in Spanish reading programs.

Correlations Between Deviates

Given that the correlations between deviates within each instructional year are quite similar for the English and Spanish data, these will be discussed together. The inter-correlations between the nine English deviates within each instructional year for the bilingual sample are presented in the four panels of Table 3; those for the Spanish deviates are displayed in Table 4. Given the linear relationship between the raw scale critical indices and the deviations of these critical indices from the aggregate growth track, the correlation pattern for the two types of measures within instructional year should be identical. The correlations between scale critical indices by instructional year for both the English and Spanish data for the bilingual sample were presented in Table 37 of Volume 5. The slight discrepancies between the corresponding tables are due to the deletion of a single (different) IRAS component for five students which showed extreme negative growth patterns. These students came from the nonborder sites where growth functions were based on only two data points and the extreme growth functions derived from them seemed untrustworthy.

As seen in the tables, the correlation coefficients are generally larger in English than in Spanish, and both show steady increases in magnitude over the instructional years. The patterns of relations, however, remain fairly stable over instructional years, and are quite similar for both English and Spanish.

The highest correlations for the formal language scales (Vocabulary Definition, Narrative Listening Comprehension, and Expository Listening Comprehension) are those obtained between each other. To a lesser degree, these scales are also related to the reading scales (Sentence Reading, Narrative Reading Comprehension, and Expository Reading Comprehension), where the relative magnitudes of the relations increase with each instructional year. Their weakest relations are with the decoding scales (Vocabulary Decoding, Letter-sound Decoding, and Letter-sound Spelling), which also grow in relative strength over the instructional years.

Similarly, for the decoding scales, the highest correlations at each instructional year are the intra-scale correlations between them. To a lesser degree, these scales are related to the reading scales, and also show an increase in relative magnitude over instructional years. Their weakest relations are with the formal language scales, again showing increases in relative magnitude with increases in schooling.

The reading scales also follow this pattern: highest relations between each other, weaker relations with decoding, and weakest relations with formal language, with the latter two sets of relations

Table 3

Interactive Reading Assessment System - English:
Correlations between Deviates by Instructional Year
for the Bilingual Sample

		SCALE								
YEAR	SCALE	VCD	VDF	LDC	LSP	SRD	NRC	ERC	NLC	ELC
	VCD	—	.52	.74	.74	.86	.83	.78	.34	.47
	VDF		—	.38	.38	.43	.43	.34	.50	.50
	LDC			—	.78	.69	.60	.59	.38	.44
	LSP				—	.64	.61	.60	.31	.38
1	SRD					—	.88	.79	.36	.47
	NRC						—	.91	.37	.48
	ERC							—	.36	.49
	NLC								—	.80
	ELC									—
	VCD	—	.75	.76	.77	.78	.77	.73	.54	.63
	VDF		—	.58	.59	.60	.62	.58	.61	.64
	LDC			—	.72	.66	.64	.60	.43	.51
	LSP				—	.63	.64	.60	.45	.51
2	SRD					—	.86	.83	.54	.66
	NRC						—	.95	.64	.77
	ERC							—	.61	.75
	NLC								—	.86
	ELC									—
	VCD	—	.77	.75	.74	.77	.77	.76	.64	.68
	VDF		—	.53	.62	.59	.62	.59	.63	.64
	LDC			—	.70	.69	.68	.67	.48	.58
	LSP				—	.64	.60	.62	.52	.56
3	SRD					—	.87	.88	.69	.72
	NRC						—	.95	.78	.81
	ERC							—	.75	.81
	NLC								—	.92
	ELC									—
	VCD	—	.81	.82	.79	.73	.88	.87	.71	.77
	VDF		—	.63	.63	.61	.78	.78	.71	.70
	LDC			—	.70	.61	.81	.78	.71	.78
	LSP				—	.53	.67	.66	.51	.55
4	SRD					—	.74	.71	.61	.64
	NRC						—	.94	.84	.86
	ERC							—	.79	.85
	NLC								—	.92
	ELC									—

Table 4

Interactive Reading Assessment System - Spanish:
Correlations between Deviates by Instructional Year
for the Bilingual Sample

YEAR	SCALE	VCD	VDF	LDC	LSP	SRD	NRC	ERC	NLC	ELC
1	VCD	-	.65	.83	.79	.83	.63	.44	.34	.39
	VDF		-	.64	.60	.62	.41	.25	.59	.57
	LDC			-	.82	.76	.53	.35	.39	.47
	LSP				-	.74	.54	.38	.37	.44
	SRD					-	.57	.37	.39	.39
	NRC						-	.69	.24	.33
	ERC							-	.15	.23
	NLC								-	.88
	ELC									-
2	VCD	-	.68	.85	.87	.86	.77	.74	.54	.53
	VDF		-	.58	.63	.61	.49	.44	.65	.60
	LDC			-	.89	.77	.64	.57	.54	.51
	LSP				-	.78	.67	.59	.53	.49
	SRD					-	.82	.78	.51	.53
	NRC						-	.96	.49	.53
	ERC							-	.44	.51
	NLC								-	.88
	ELC									-
3	VCD	-	.72	.82	.83	.80	.76	.77	.35	.47
	VDF		-	.57	.59	.56	.58	.56	.41	.48
	LDC			-	.82	.69	.64	.61	.30	.37
	LSP				-	.74	.64	.62	.33	.39
	SRD					-	.80	.79	.42	.50
	NRC						-	.96	.49	.57
	ERC							-	.48	.58
	NLC								-	.84
	ELC									-
4	VCD	-	.87	.82	.89	.73	.75	.78	.51	.56
	VDF		-	.60	.79	.55	.62	.66	.50	.56
	LDC			-	.79	.73	.71	.69	.36	.40
	LSP				-	.71	.74	.75	.47	.48
	SRD					-	.80	.78	.34	.42
	NRC						-	.97	.57	.61
	ERC							-	.60	.66
	NLC								-	.86
	ELC									-

increasing in relative magnitude with increases in instructional years.

The concept of separable components in reading is a question of substantial importance both theoretically and practically. Many experts would probably hold either that (a) reading is best described as a one-dimensional aptitude or achievement, or (b) reading is a complexly related set of interactive dimensions. Both of these positions lead to the same conclusion -- no discernible structure in the relations among component reading skills. The present findings reveal a structure, which in first grade comprises (a) formal skills in processing spoken language, (b) level of skill in decoding single words, and (c) ability to comprehend printed materials. This structure becomes less clearcut in later grades, with high correlations among all of the IRAS components.

At least three hypotheses can be put forward to explain this trend. First, it may be a fundamental developmental tendency; with increasing age, competence tends to converge to a level reflective of the domain and the individual. Second, it may be an inherent demand of the curriculum. In the beginning a student may be more or less talented in certain aspects of reading, but in the later grades the requirements of skilled reading require the individual to "pull it all together," a task that is realized with more or less success by a given individual. Finally, it may be that the convergence is a product of the way that reading instruction is organized. If a student is identified as wanting or with promise in a particular area of reading, or "in general," then the level of instruction will be designed to suit that perception.

Predictor Variables

In this section, the two classes of predictor variables employed in the integrative analyses are discussed giving their associated descriptive statistics first, then their inter-correlations (i.e., an assessment of the degree of predictor collinearity). The first class of indices contains the precursor variables, and includes (a) English and Spanish oral language skill at kindergarten entry and (b) knowledge of the English and Spanish alphabet (for Instructional Year 1 analyses), or previous year's performance (for analyses of Instructional Years 2 through 4). Instruction variables constitute the second class of indices, including (a) the number of years of Spanish reading, (b) the dimensions of observed instruction, (c) the dimensions of planned instruction, (d) the average percentage of days present in school, and (e) the three orthogonal site contrasts.

Descriptive Statistics

Table 5 presents the descriptive statistics for each of these predictors for each instructional year for the bilingual sample. The variables are listed along the left-hand margin with instructional years defining the right-hand columns. For each scale, the mean (M)

Table 5

Descriptive Statistics on Predictors by Instructional Year for the Bilingual Sample

Scale	Statistic	Instructional Year (Number of Cases) 1 (249)	2 (248)	3 (93)	4 (58)
ENGCATG	M	0.09	0.08	-0.03	0.07
	S	1.00	1.00	1.01	1.01
SPNCATG	M	0.10	0.11	0.56	0.52
	S	1.00	1.00	0.83	0.86
SFSTALPH ENGLISH	M	-0.36	-	-	-
	S	0.94	-	-	-
SFSTALPH SPANISH	M	-0.59	-	-	-
	S	0.81	-	-	-
PROGGRP	M	1.42	1.44	2.12	2.16
	S	1.74	1.73	2.10	2.30
PROGY	M	-0.29	-0.32	-0.23	-0.45
	S	0.96	0.95	0.98	0.70
RAMOS-M ENGLISH	M	0.15	0.12	0.05	0.03
	S	0.36	0.33	0.23	0.18
CHECK-M ENGLISH	M	0.20	0.11	0.03	0.02
	S	0.40	0.31	0.18	0.13
RAMOS-M SPANISH	M	0.71	0.76	0.81	0.88
	S	0.45	0.43	0.39	0.33
CHECK-M SPANISH	M	0.70	0.65	0.54	0.69
	S	0.46	0.48	0.50	0.47
PRESPRY	M	93.52	94.32	94.50	95.51
	S	6.60	7.48	7.37	7.21
SITE05	M	0.00	0.00	-0.01	-0.01
	S	0.01	0.01	0.00	0.00
SITE35	M	-0.00	0.00	-	-
	S	0.01	0.01	-	-
SITE12	M	0.00	0.00	0.00	-0.00
	S	0.01	0.01	0.02	0.02

and standard deviation (S) are listed. Scale names are mnemonic, as specified below:

- ENGCATG: English kindergarten entry category based on teacher ratings (dichotomous)
- SPNCATG: Spanish kindergarten entry category based on teacher ratings (dichotomous)
- SFSTALPH: Stanford Foundation Skills Test alphabet production/recognition category, separately for English and Spanish (dichotomous)
- PROGGRP: Number of years enrolled in Spanish reading (continuous)
- PROGYx: Indicator of enrollment in Spanish reading in Instructional Year *x* (dichotomous)
- RAMOS-M: Indicator of missing RAMOS factor data in the present year (dichotomous)
- CHECK-M: Indicator of missing Checklist factor data in the present year (dichotomous)
- PRESPRYx: Average percentage of days attended in Instructional Year *x* (continuous)
- SITE05: Site contrast between border and nonborder sites (dichotomous)
- SITE35: Site contrast between nonborder sites (dichotomous)
- SITE12: Site contrast between border sites based on degree of urban influence (dichotomous)

In addition to these predictors, the RAMOS and Checklist factor scores were also included in the integrative analyses. The descriptive data for these are presented in Table 6 (RAMOS data) and Table 7 (Checklist data) for the English dimensions of instruction, and Tables 8 and 9 for the Spanish instructional dimensions. A discussion of each of these variables follows.

Precursors

The index of oral language skill at kindergarten entry was based on an approximate median split of the teacher rating data from the Oral Language Proficiency Rating Scale (OLPRS). The linear functions derived from these data were presented in Table 8 of Volume 4. The median split was based on the Fall kindergarten ratings divided into two categories to achieve an approximate even distribution of students within the two languages. Based on the five-point OLPRS rating scale, a value of 3.0 or above defined the high English category, and 4.0 or above defined the high Spanish category. In both languages, the high category was coded as +1 and the low category as -1. As such, it is important to remember that aggregate skill in English at kindergarten entry (as rated by teachers) was lower than that for Spanish, and the resulting definitions of relatively "high" and "low" skill based on the two median splits reflect this difference.

From Table 5, the summary values for English entry indicate approximately equal numbers of high (3.0 or above) and low entry target students, with slightly more high entry students in

Table 6

Reading and Mathematics Observation System - English:
Descriptive Statistics on Factor Scores for Each Instructional Year
for the Bilingual Sample

FACTOR	STATISTIC	INSTRUCTIONAL YEAR 0	1	2	3	4
ETT	M	0.009	-0.001	-0.003	0.000	0.000
	SD	0.349	0.338	0.348	0.349	0.302
DGI	M	-0.007	0.000	-0.002	0.000	0.000
	SD	0.322	0.361	0.334	0.370	0.333
QFL	M	0.000	-0.013	-0.002	-0.002	-0.002
	SD	0.366	0.434	0.380	0.427	0.350
ADC	M	0.060	0.022	0.009	0.000	0.005
	SD	0.568	0.645	0.619	0.618	0.625
PRD	M	-0.016	0.000	-0.002	0.000	-0.001
	SD	0.377	0.353	0.329	0.274	0.377
SMT	M	-0.032	0.009	0.009	0.000	0.000
	SD	0.499	0.442	0.490	0.367	0.265
NST	M	0.000	0.000	0.000	0.000	0.000
	SD	1.000	1.000	1.000	1.000	1.000
	N	140	215	224	94	59

Table 7

Reading and Mathematics Observation System - Spanish:
Descriptive Statistics on Factor Scores for Each Instructional Year
for the Bilingual Sample

FACTOR	STATISTIC	INSTRUCTIONAL YEAR 0	1	2	3	4
QFL	M	0.005	-0.012	-0.001	0.000	0.000
	SD	0.309	0.391	0.376	0.240	0.117
DGI	M	0.000	-0.012	0.000	0.000	0.000
	SD	0.420	0.530	0.433	0.547	0.105
ETT	M	0.001	0.006	0.000	0.000	0.000
	SD	0.451	0.566	0.523	0.375	0.545
NST	M	0.000	-0.008	-0.004	0.000	0.000
	SD	0.471	0.455	0.534	0.486	0.587
ADC	M	0.000	0.000	0.000	0.000	0.000
	SD	0.724	0.783	0.729	0.761	0.667
SMT	M	0.000	0.000	0.000	0.000	0.000
	SD	0.460	0.497	0.461	0.416	0.121
CNT	M	-0.005	-0.003	0.014	0.010	0.000
	SD	0.413	0.392	0.363	0.254	0.285
	N	67	73	62	20	9

869

Table 8

Checklist - English:
Descriptive Statistics on Factor Scores for Each Instructional Year
for the Bilingual Sample

FACTOR	STATISTIC	INSTRUCTIONAL YEAR 0	1	2	3	4
ACM	M	0.010	0.040	0.000	0.000	0.000
	SD	0.648	0.732	0.578	0.580	0.570
QFL	M	0.004	-0.002	0.001	0.000	0.000
	SD	0.337	0.323	0.340	0.295	0.280
STW	M	0.000	0.000	0.000	0.000	0.000
	SD	0.731	0.807	0.810	0.806	0.780
PMT	M	-0.011	-0.010	-0.004	-0.001	0.000
	SD	0.473	0.531	0.433	0.322	0.381
GRV	M	0.000	0.004	0.006	0.002	0.000
	SD	0.487	0.447	0.423	0.482	0.445
	N	163	203	228	97	60

Table 9

Checklist - Spanish:
Descriptive Statistics on Factor Scores for Each Instructional Year for the Bilingual Sample

FACTOR	STATISTIC	INSTRUCTIONAL YEAR 0	1	2	3	4
ADC	M	0.000	0.000	0.000	0.000	0.000
	SD	0.596	0.578	0.520	0.533	0.540
STW	M	0.000	0.000	0.000	0.000	0.000
	SD	0.564	0.682	0.784	0.804	0.679
PMT	M	0.000	0.000	0.000	0.000	0.000
	SD	0.389	0.421	0.419	0.506	0.246
DTC	M	0.004	0.000	0.000	0.000	0.008
	SD	0.652	0.603	0.567	0.504	0.380
NST	M	0.000	0.000	0.000	0.000	0.000
	SD	1.000	1.000	1.000	1.000	1.000
	N	93	75	90	48	21

871

Instructional Years 1, 2 and 3, and slightly more low entry students in Instructional Year 4. For Spanish entry skill, Instructional Years 1 and 2 similarly contain slightly more high entry target students (4.0 or above), but Instructional Years 3 and 4 show substantially more such target students (a ratio slightly above 3 to 1). Such would be expected given the cohort structure of the study, as students tracked through third and fourth grade came entirely from the border sites where support for Spanish and thus, resulting skill, was much greater.

Two indices of previous year's achievement were employed. For analyses of the first instructional year, the dichotomous contrast for knowledge of the alphabet was used (English or Spanish depending on the IRAS data set analyzed). In subsequent instructional years, the actual scale value from the previous instructional year was employed. The alphabet knowledge index was derived from data on the combined recognition and production tasks of the <u>Stanford Foundation Skill Test</u>. The summary data for these indices were treated in Volume 5, with the English data presented in Tables 4 and 5, and the Spanish data in Tables 11 and 12. For the current analyses, approximate median splits were generated from these data. For English, the relatively high alphabet knowledge category was defined by averaged recognition/production scores greater than 50%, and for Spanish, the high alphabet knowledge category was defined by averaged recognition/production scores greater than 20%. Again, within both languages, the high category received a +1 value, the low category, a value of -1. As with oral language entry skill, it is important to remember the English-Spanish difference: English alphabet knowledge (at kindergarten) was greater than knowledge of the Spanish alphabet, and the resulting definitions of relatively "high" and "low" skill reflect this dfference in the two median splits.

As seen in Table 5, the values for both the English and Spanish alphabet knowledge indices indicate more low knowledge target students than high knowledge target students (a 2 to 1 ratio in English and a 3 to 1 ratio in Spanish).

<u>Instruction</u>

Two indices reflecting enrollment in Spanish reading programs were computed. These indices were based on the instructional program variable discussed in Volume 6, with descriptive statistics for the variable presented in Table 50 of that volume. The first index captured the students' history of Spanish reading. For this variable, no Spanish reading was coded as 0; Spanish reading only in kindergarten was coded as 1; Spanish reading in K and 1 was coded as 2; K-2 as 3; K-3 as 4; and K-4 as 5. In addition, a second index of Spanish reading was also computed. This variable simply coded whether or not a student was currently enrolled in Spanish reading for the instructional year under analysis, with a value of +1 indicating such enrollment, and -1 indicating no Spanish reading enrollment for the relevant instructional year.

The data from the program history variable (PROGGRP) given in Table 5 show the expected increase for the number of years of Spanish reading in both mean and standard deviation over instructional years. For the first two instructional years, the average number of years of Spanish reading is about 1.4, increasing to 2.2 by the last two instructional years.

Two sources of specific instructional dimensions were employed in the integration analyses, those based on observed instruction (derived from the Reading and Mathematics Observation System), and those based on planned instruction (derived from the Teacher Checklist). Summary descriptions of these data were fully given in Volume 6 as follows: (a) English observed instruction factor scores in Tables 8 and 9, (b) English planned instruction factor scores in Tables 30 and 31, (c) Spanish observed instruction factor scores in Tables 16 and 17, and (d) Spanish planned instruction factor scores in Tables 38 and 39. As an aid to the reader, the overall descriptive statistics for these factors for each language and instrument by instructional year have been reproduced here in Tables 6 through 9.

The descriptive statistics for the instruction factor scores were discussed in detail in Volume 6, and will only be briefly reviewed here. In general, the mean values for each factor within each instructional year are close to 0. Recall that the procedure for deriving factor scores involved standardization (by instructional year) of the averaged summary indices first, then weighting and summing of factor component values. As such, the resulting values are not themselves standardized, and thus standard deviations of 1 are not expected, nor found, as can be seen from the tables.

An index of whether the instructional indices just described contained any estimated values due to missing data was computed separately for observed instruction and planned instruction. Such estimates were used in order to maximize the number of cases available for analysis. For each language and instrument, these values were computed by noting whether any of the instructional factor scores were estimated (using the mean value for the appropriate factor and instructional year) due to missing data, with values of +1 indicating that at least one of the factor scores was estimated, and a value of 0 indicating that none of the factor scores were so estimated.

From Table 5, the amount of estimated English RAMOS data was generally low, and declined over the four instructional years from 15% of the cases to 3%. Estimated English Checklist data also showed a similar trend, declining from 20% to 2%. These missing cases in large degree reflect the number of students who were enrolled in exclusive Spanish reading programs and thus did not receive English reading. Note that the estimates, being mean values, do not enter into the evaluation of the contribution of individual instructional factors to reading deviate predictions. They simply maintain the maximum number of cases (for purposes of assessing the effects of the other predictor variables), allowing the instructional assessments to be based on the actual data values available.

873

For the Spanish instructional factor estimates, the values are much higher, as would be expected given the number of students who did not receive Spanish reading instruction. For the RAMOS factors, the percentage of estimated cases is fairly stable across instructional years, averaging about 80%. Similarly, for the Checklist data, within each instructional year about 65% of the cases contained missing data. Given the relatively large numbers of estimated values in Spanish reading instruction, one can expect an artifactual reduction in the predictive power of these independent varibles.

Given the emphasis on "time on task", the integrative analyses contained an index of the average percentage of days present in school for the instru[illegible]onal year under analysis. These data were based on actual attendance records collected at the end of each academic year for each of the target students. As seen in Table 5, the average percentage is high at about 95%, but the standard deviation value of about 7% indicates that some students were absent from school for a substantial amount of time.

Finally, to assess any remaining contributions of instruction associated with specific sites, the integrative analyses employed three orthogonal site contrast indices. The first variable contrasted the border and nonborder sites. The second site contrast allowed the two nonborder sites to be compared. The third index defined a contrast <u>within</u> the border sites which allowed the two sites with relatively <u>low</u> degrees of urban influence (Sites 0 and 1) to be compared with the remaining border site which showed a relatively high degree of urban influence (Site 2). These site differences were discussed in the site descriptions presented in Appendix A of Volume 2. Orthogonal contrast codings were used by employing weights based upon the number of target students at each site. Thus, for the border versus nonborder site contrast (Sites 0, 1, and 2 versus Sites 3 and 5), border site target students received a coding of -1/108, while nonborder site target students received values of +1/146, reflecting the number of target students found within each site category. For the contrast within the nonborder sites, Site 3 students received a value of -1/[illegible]5, Site [illegible] [illegible]get students received a value of +1/70, and the border site stu[illegible] [illegible]ites 0, 1, and 2) received values of 0. Finally, for the wi[illegible] [illegible]order site contrast, target students from Sites 0 and 1 received values of -1/72, Site 2 target students received values of +1/36, and the nonborder site target students (Sites 3 and 5) received values of 0.

Having given the descriptive statistics for the set of predictor variables, we next turn to a discussion of their interrelatedness.

Collinearities Among Predictors

As noted earlier in the volume, one of the major threats to a clearly interpretable regression analysis is the presence of confoundings or collinearities in the set of predictor factors. This section reviews the correlations among the predictors for each of the dependent measures.

The correlation matrices in this section all have the same general form, paralleling the order in which predictors were introduced into the regression equation. Each matrix contains entries both above and below the main diagonal; those above are for Year 1 data, and those below the diagonal are for Year 2 data. The data for Years 3 and 4 were examined, but are not reported for the following reasons. First, these data are from restricted cohorts, and so are of somewhat less interest than those from the first two years. Second, the basic patterns and conclusions are unchanged between the first two and last two years.

Most of the predictors form an identical set over the set of outcome measures. That is, the program, instructional, and site predictors are largely the same as one goes from Vocabulary Decoding to Vocabulary Definition to Letter-sound Decoding, and so on. The precursor predictors (previous year's performance) do change, of course. Moreover, there are some changes in the nature of the sample from one measure to another, and these have both potential and real effects on certain features of the collinearity structure. Three measures are most effected -- Sentence Reading, Expository Reading Comprehension, and Expository Listening Comprehension -- because they were introduced as part of the test battery upon entry of the third cohort of students into first grade. In any event, it seemed worthwhile to present the collinearity matrices for each of the outcome measures, in order to determine any variations that might exist.

Predictors for English Reading

Table 10 displays the collinearities for Years 1 and 2 for English Vocabulary Decoding. We will examine this table in some detail, as a prototype of those that follow.

Precursor factors. The first predictor, row-wise and column-wise, is entry language classification, the student's oral language competence as rated by the kindergarten teacher on entry to kindergarten, divided into above and below median categories as a contrast. In Year 1, this measure for Vocabulary Decoding is correlated .32 with performance at the beginning of Year 1 (actually the Alphabet Knowledge contrast from the beginning of kindergarten or first grade). The corresponding correlation for Year 2 is .39; here the previous year's performance is actually the deviation for Vocabulary Decoding as measured at the end of Year 1.

Entry language classification is negatively correlated with the two measures of program: PROGGRP is the number of years that a student received Spanish reading instruction, and PROGYn is an index that is 1 if the student was assigned to Spanish instruction during the target year and 0 otherwise. In both years the correlation ranges around a value of -.4 for both program indices; the student who is rated relatively high in English on kindergarten entry is less likely to be assigned to Spanish reading instruction, a validation of sorts that children are not assigned to these bilingual programs at random.

Table 10

Correlation Matrix Showing Extent and Pattern of Collinearities
Among Predictors for VOCABULARY DECODING (IRAS -E)
Year 1 above diagonal; Year 2 below diagonal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 ENGCATG	100	32	-38	-43	9	1	13	1	-2	-11	0	-17	22	-1	28	4	-13	-25	3	6	-10	-4
2 PREVYR	39	100	-27	-28	0	7	-2	2	2	6	7	8	13	17	13	-9	-8	-22	8	-17	1	9
3 PROGGRP	-37	-28	100	94	-32	30	8	21	24	29	-5	26	-18	-2	-2	3	26	34	-3	-37	25	-14
4 PROGYx	-37	-32	93	100	-35	24	-5	19	21	33	1	24	-19	2	-7	2	34	51	-4	-30	26	-10
5 R-ETT	4	6	-21	-23	100	-2	8	-20	-20	-11	-12	-14	-17	-13	-15	5	0	-18	-8	21	-46	7
6 R-DGI	4	-3	29	24	-18	100	-9	47	56	16	10	10	5	-19	18	1	1	-2	-5	-52	4	17
7 R-QFL	6	-5	-5	-11	36	33	100	-6	3	-26	-36	-11	6	-25	19	21	1	-23	0	21	6	-42
8 R-ADC	-3	0	0	2	0	-4	15	100	34	30	11	1	17	-9	29	-1	-1	11	-8	-33	34	4
9 R-PRD	11	29	21	15	-28	37	-13	-2	100	22	-4	5	5	-22	33	-6	0	1	4	-51	18	0
10 R-SMT	-2	12	17	11	-33	-18	-37	-26	15	100	6	-3	-3	-12	24	-3	0	15	6	-42	25	-15
11 R-NST	9	-7	-34	-27	11	-25	18	1	-49	-10	100	8	15	38	-4	-21	-1	20	-1	-5	16	63
12 C-ACH	24	48	-16	-14	1	-20	-7	16	27	-2	-13	100	3	52	-41	-61	35	0	-1	-27	9	48
13 C-QFL	20	23	-25	-24	20	-27	1	14	-4	0	20	23	100	0	33	2	-6	0	-3	3	21	4
14 C-STW	6	10	-35	-32	19	-25	2	-2	-31	-1	19	16	1	100	-35	-47	25	-1	2	-2	17	54
15 C-PMT	8	18	26	15	-14	34	7	-2	42	16	-34	-10	14	-36	100	6	-23	-1	19	-20	26	-41
16 C-GRV	-8	-14	-3	0	13	-3	2	-8	-14	-1	32	-53	6	1	-11	100	-17	2	7	20	-10	-48
17 R-MISS	-28	-22	31	41	-1	0	0	0	0	0	0	1	4	-9	3	0	100	54	-13	2	21	10
18 C-MISS	-24	-26	29	42	-10	-6	-9	-2	-7	6	1	0	0	0	1	0	68	100	-13	12	32	-1
19 PRESPRY	0	9	-15	-10	17	-15	5	8	-1	-7	13	8	7	9	-4	4	-17	-7	100	1	18	3
20 SITE05	6	-13	-35	-26	46	-21	19	2	-46	-52	42	0	25	18	-33	17	0	13	12	100	-1	-2
21 SITE35	-8	16	25	27	-2	-9	-33	-16	13	36	-20	7	13	7	11	9	28	32	-5	1	100	0
22 SITE12	-3	-8	-13	-15	15	-24	3	-9	-30	14	28	9	3	52	-31	0	-10	-5	18	-1	0	100

Next in the list of predictors is previous year's performance. This measure is negatively correlated with the program variables (PROGGRP and PROGYn) in both years at a level of about -.3. That is, a student whose previous year's performance is below par (in English Vocabulary Decoding) is more likely to have been assigned to relatively more years of Spanish reading instruction, and to be in a Spanish reading instruction program during the target year.

The only other consistent pattern in which PREVYR is related to other predictors is in Year 2, where the correlation with the Checklist ACM factor is positive, a relation that shows up for several other IRAS measures. ACM is the Checklist factor measuring the reported emphasis on comprehension compared with decoding; students whose previous year's Vocabulary Decoding scores were relatively positive were more likely to receive relatively greater emphasis on comprehension activities. This pattern reappears in later tables in this series for IRAS measures tapping decoding skill, with the exception of Expository Reading Comprehension, but is not found for the measures of listening comprehension. The finding has a straightforward interpretation: students whose decoding skills are reasonably well developed are more likely to receive more instruction in comprehension. The down-side is that students who are experiencing difficulty with decoding are more likely to be kept at the level of a phonics emphasis.

Program predictors. Next we consider the program predictors. For Years 1 and 2, the two contrasts are highly correlated (>.90), not just for Vocabulary Decoding but for all of the IRAS measures, in both Spanish and English. The correlations drop somewhat in Years 3 and 4, but remain relatively high (.8 in Year 3 and .7 in Year 4). The relation is partly a student-specific effect, partly a site-specific effect. The student-specific effect, mentioned above, relates to the student's entry language and progress in English reading.

The site-specific effect is understandable in terms of program needs and resources. Certain sites, mostly those at the border, are more likely than others to offer Spanish reading instruction throughout the primary grades. Certain students, mostly those whose oral English is poorly developed, are likely to be assigned to Spanish reading instruction and to remain in that assignment during most of the primary years. These effects can be seen in the negative correlations with the SITE05 contrast, the positive correlation with SITE35, and the negative correlation with ENGCAT. Border sites received a "-" contrast value in SITE05; students at these sites were more likely to receive more years of Spanish reading instruction, leading to a negative correlation. Similarly, Site 3, assigned a "-" value in the SITE35 contrast, was a relatively less urban locale than Site 5, and was less likely to provide Spanish reading instruction to its students.

Because orthogonal contrasts are not widely used in regression analyses, it may be of value to expand slightly on this example. As noted, in the SITE05 contrast nonborder sites were assigned a positive weighting and border sites a negative weighting. The assignment is completely arbitrary, and affects only the sign of the correlation. The effect of the contrast is to identify the difference between the two types of sites. A negative correlation between the site contrast and "years in a Spanish reading instruction program" is equivalent to the two-by-two table presented below:

Years in Spanish Reading Instruction

Site	Few ("-")	Many ("+")
Border ("-")	Not many cases	Lots of cases
Nonborder ("+")	Lots of cases	Not many cases

Both the site contrast and the program variables are categorical, and so for practical purposes the correlation is a convenient means of expressing the value of a contingency analysis.

Another source of collinearity with the program variables arises in association with the MISSING indicators for the RAMOS and the Checklist factors. Recall that whenever one of the instructional factor scores was missing, the MISSING indicator was set to +1, and the mean value of the factor was assigned to the missing slot. This procedure kept a maximum number of observations in the regression analyses, but also allowed us to monitor the potential impact of this strategy. As one would expect, the MISSING indicators for RAMOS and Checklist are correlated, generally within the range of .5 to .7; teachers who do not plan instruction in Spanish or English, respectively, are more likely to be found not to be teaching in that language.

The relations between PROG and MISSING in both English and Spanish are also as expected. One of the most common reasons for a missing RAMOS or Checklist in English occurred on those occasions when the entire lesson was in Spanish, and contrariwise. Accordingly, English MISSING was more likely to be incremented for students who were assigned to one or more years of Spanish instruction (a positive correlation, as observed), and Spanish MISSING was less likely to incremented for students assigned to Spanish reading instruction (a negative correlation, which will be seen in later tables). Since the entry language classification rating was related to program assignment, it was also correlated with MISSING, more noticeably for Spanish than for English.

The overall program correlation pattern, as a consequence, takes the following form. Some students are more likely than others to enter school with relatively limited skills in spoken English. Such students are more typical at some of the sites than others; those sites with higher proportions of Limited English Proficient students were also more likely in this study to have access to bilingual teachers. At these sites, Limited English Proficient students were likely to be assigned to Spanish reading instruction, and to remain in that program for two or three years. As it turns out, we could have used only a single program factor; including both PROGGRP and PROGYn weakened the analysis somewhat because of the redundancy of the two factors. The collinearity of these factors with the precursor factors (language and previous achievement) and with the missing instruction data indicators is a separate matter; untangling the relation by nested contrasts would provide a clearer picture of program effects, in our opinion.

Instructional predictors. Turning next to the instructional factors, we note that the predictors from RAMOS and Checklist were constructed to achieve a high degree of orthogonality through a series of factor analyses. On the other hand, these analyses were conducted over all four years of data, which means that there is no guarantee of orthogonality for any given year. Examination of the collinearity patterns reveals that systematic relations between the factor scores are of limited extent, and further, factor scores do not appear to be systematically related to site or program variables in most instances. A few exceptions will be described in this section.

Taking a correlation of .3 as a touchstone, 24 of the 132 correlations (18%) between the RAMOS and Checklist predictors equal or exceed that criterion. Nine of these are .4 or greater. One pair is large in both Years 1 and 2: Amount of Comprehension (ACM) instruction and Group Vocabulary Instruction (GRV) are positively correlated in excess of .5 in both years. Otherwise no consistent patterns emerge in the submatrix of instructional factor scores.

Instruction at different sites. Certain features of the instructional configuration are systematically related to site contrasts in both years. PRD and SMT are linked to SITE05, while STW and PMT are related to SITE12. PRD is the observational variable measuring rated productivity, and SMT assesses the availability of secondary materials. For both, higher values are associated with what are generally construed to be more positive instructional conditions. These are positively correlated with the border-nonborder site contrast, which in this instance is positive for the nonborder sites. The implication is that in the first two years of the study, high levels of productivity and a greater diversity of materials were typical of the nonborder sites, compared with the border sites. This relation was based on observational summaries of classroom instruction.

Parenthetically, the contrast between the border and nonborder sites is associated with two other observational predictors in a pattern that is not totally consistent over years, but such that the direction of the relation is consistent. Briefly, in the nonborder sites, students were observed to spend relatively more time engaged with text in both first and second grades, and relatively less time in whole-group teacher-directed instruction (i.e., more time with small-group activities and some parallel independent work).

STW is an index of the teacher's report about the amount of seatwork to be assigned, while PMT is the quality of the primary materials as designated by the teacher. The association of these factors was with the contrast between the border sites, one of which was relatively closer to a city, the other two of which were relatively isolated. This confounding, which is probably idiosyncratic to policies at the sites, can be described as follows. The more heavily urban influenced site, in contrast with the two more isolated sites, relied more on worksheets and seatwork, and less on textbooks and other relatively more demanding primary text materials. This may be explained in part by the nature of the reading program at this site. As reported in Volume ? this site provided individualized reading instruction managed th ugh the use of student contracts in Instructional Years 1 and 2. As a consequence, reliance on worksheets and seatwork assignments was more prevalent in classrooms at this site than in the other two border sites where reading instruction tended to conform more closely to traditional small-group instruction built around a basal reader.

The contrast for the degree of urban influence among the border sites is also associated with four other instructional predictors, but for Year 1 only. The correlations with QFL, NST, ACM, and GRV, all greater in absolute value than .4, yield the following configuration for first graders: In the more heavily urban influenced site, as contrasted with the two more isolated sites, first grade included less emphasis on formal language during reading, a larger number of students were observed in reading groups, and there was less time spent on decoding and group instruction on word meaning. Such differences may be expected, give the characteristics of the individualized reading instructional program at this site and the district's policy of delaying phonics instruction until the student had completed the primer level of the assigned primary textbook.

Having gone through Table 10 in some detail, we can dispense with the remainder of Tables 11 to 18 in relatively short order. As noted at the beginning of the section, collinearities with entry language classification and previous year's achievement deserve some attention, and certain IRAS measures were not introduced until the third cohort. First, let us consider collinearities with entry language category and previous year's achievement. This relation is virtually unchanged over measures in Year 1, because both indices are the same (recall that Alphabet Knowledge is the measure of previous achievement), and variations arise only because of slight changes in the sample. For Year 2, the change in the strength of the correlation over measures has an interesting structure. In general, the relation of the original rating of oral language on entry to kindergarten is strongest with IRAS measures of oral language -- two years later. The correlations with Vocabulary Definition and Narrative Listening Comprehension are both .46. Next highest are correlations with Vocabulary Decoding and Narrative Reading Comprehension, both slightly

Table 11

Correlation Matrix Showing Extent and Pattern of Collinearities Among Predictors for VOCABULARY DEFINITION (IRAS - E)
Year 1 above diagonal; Year 2 below diagonal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 ENGCATG	100	31	-38	-44	9	1	13	1	-2	-11	0	-17	22	-1	29	4	-14	-26	3	7	-10	-3
2 PREVYR	47	100	-27	-29	0	7	-2	2	2	6	8	8	13	17	14	-9	-9	-23	8	-16	1	10
3 PROGGRP	-36	-46	100	94		30	8	21	25	29	-5	26	-18	-2	-2	3	26	34	-3	-36	25	-13
4 PROGYx	-37	-45	93	100		24	-5	19	21	33	1	24	-19	2	-7	2	32	50	-4	-30	26	-10
5 R-ETT	4	6	-21	-23	30	-2	8	-20	-20	-11	-22	-14	-17	-13	-15	5	0	-18	-8	21	-46	7
6 R-DGI	4	-20	29	24	-18	100	-6	47	56	16	10	10	5	-19	18	1	1	-2	-5	-52	4	17
7 R-QFL	6	-9	-5	-11	36	33	100	-7	3	-26	-36	-10	6	-25	19	21	1	-23	0	21	6	-42
8 R-ADC	-3	8	-1	2	0	-4	15	100	34	30	11	1	17	-9	29	-1	1	11	-8	-33	34	4
9 R-PRD	11	12	22	15	-28	37	-13	-2	100	22	-4	5	5	-22	33	-6	0	1	4	-51	18	0
10 R-SMT	-2	-4	17	11	-33	-18	-37	-26	15	100	6	-3	-3	-12	24	-3	0	15	6	-4	25	-15
11 R-NST	10	7	-34	-28	11	-25	18	1	-49	-10	100	8	15	38	-4	-21	-1	21	-1	-5	16	63
12 C-ACM	24	29	-16	-14	1	-20	-7	16	27	-2	-13	100	3	52	-41	-61	36	0	-1	-27		49
13 C-QFL	21	24	-25	-24	20	-27	2	14	-4	0	20	23	100	0	33	2	-6	0	-3	3		4
14 C-STW	6	12	-35	-32	19	-25	2	-2	-31	-1	19	16	1	100	-35	-47	25	-1	2	-2		55
15 C-PMT	8	-5	26	15	-14	34	8	-2	42	16	-34	-10	14	-38	100	6	-24	-1	19	-20	26	-41
16 C-GRV	-9	-7	-4	0	13	-3	2	-8	-14	-1	32	-53	7	1	-11	100	-17	2	-7	20	-10	-48
17 R-MISS	-26	-29	30	39	-2	0	0	0	1	C	0	1	4	-9	3	0	100	53	-13	3	21	11
18 C-MISS	-22	-29	28	40	-10	-6	-9	-2	-7	6	1	0	0	0	1	0	66	100	-13	13	33	0
19 PRESPRY	1	-4	-16	-10	18	-15	5	8	-1	-7	13	8	8	9	-4	4	-19	-8	100	1	18	3
20 SITE05	7	-4	-26	-27	46	-22	19	2	-46	-52	42	0	25	18	-33	17	-3	12	12	100	-1	-3
21 SITE35	-7	-9	24	26	-2	-9	[illegible]	-16	14	37	-20	8	13	7	11	9	26	30	-1	0	100	0
22 SITE12	-3	-8	-14	-15	15	-24	5	-10	-30	14	29	-1	3	52	-31	0	-10	-5	18	-2	0	100

882

883

Table 12

Correlation Matrix Showing Extent and Pattern of Collinearities
Among Predictors for LETTER-SOUND DECODING <IRAS - E>
Year 1 above diagonal; Year 2 below diagonal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 ENGCATG	100	31	-38	-43	11	1	13	0	-2	-10	-1	-17	21	-1	28	4	-14	-26	3	5	-11	-4
2 PREVYR	31	100	-27	-28	2	7	-3	0	2	6	7	9	11	17	12	-10	7	-21	7	-18	0	9
3 PROGGRP	-36	-17	100	94	-32	30	9	22	24	-5	26	-18	-2	-2	-2	3	28	35	-3	-37	26	-14
4 PROGYx	-37	-20	93	100	-36	24	-5	20	21	32	2	23	-18	2	-6	2	35	51	-4	-30	28	-10
5 R-EIT	4	11	-21	-23	100	-2	9	-18	-20	-12	-21	-15	-15	-13	-13	5	-1	-19	-7	22	-45	7
6 R-OGI	4	-7	29	24	-18	100	-9	47	56	16	10	10	4	-19	17	1	1	-2	-5	-52	4	17
7 R-QFL	6	-12	-6	-11	36	33	100	-7	3	-26	-37	-10	6	-26	19	21	1	-23	-1	21	5	-42
8 R-ADC	-2	-5	-2	2	1	-5	14	100	34	31	10	1	15	-9	28	-1	1	11	-9	-34	33	4
9 R-PRD	11	16	22	15	-28	37	-13	-2	100	22	-4	5	5	-22	34	-5	0	1	4	-51	18	0
10 R-SMT	-3	5	18	11	-33	-18	-37	-26	15	100	6	-3	-2	-12	25	-3	0	15	6	-42	26	-15
11 R-NST	10	3	-34	-28	12	-25	18	0	-49	-11	100	8	14	38	-5	-21	0	21	-1	-5	16	63
12 C-ACM	25	29	-16	-14	1	-20	-7	17	27	-2	-12	100	3	52	-41	-61	36	0	-1	-27	9	48
13 C-QFL	20	26	-24	-23	20	-27	2	16	-4	0	21	23	100	-1	31	2	-6	1	-4	2,	20	4
14 C-STW	5	1	-35	-32	19	-25	2	-2	-31	-1	19	16	1	100	-35	-47	25	-1	2	-2	17	54
15 C-PMT	7	7	26	15	-15	34	8	-1	42	16	-34	10	14	-37	100	6	-23	0	18	-21	25	-41
16 C-GRV	-9	5	-3	1	13	3	2	-8	-14	-2	31	-52	7	1	-11	100	-17	2	-7	20	-10	-48
17 R-MISS	-28	-16	31	41	-1	0	0	-1	1	0	1	1	4	-8	3	0	100	53	-13	2	20	10
18 C-MISS	-24	-19	29	42	-10	-6	-9	-2	-7	6	1	0	0	0	1	0	68	100	-13	12	33	-1
19 PRESPRY	0	2	-15	-10	17	-15	5	9	-1	-7	14	8	7	9	-4	5	-18	-7	100	0	17	3
20 SITE05	6	5	-35	-25	46	-21	20	3	-46	-53	43	0	12	7	10	29	33	-5	0	100	-2	-2
21 SITE35	-9	28	26	28	-3	-9	-33	-14	14	36	-20	8	12	7	10	9	29	33	-5	0	100	0
22 SITE12	-3	-21	-14	-15	15	-24	5	-10	-30	14	29	-1	3	52	-31	0	-10	-5	18	-2	0	100

Table 13

Correlation Matrix Showing Extent and Pattern of Collinearities
Among Predictors for LETTER-SOUND SPELLING (IRAS - E)
Year 1 above diagonal; Year 2 below diagonal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 ENGCATG	100	31	-38	-43	9	2	13	2	-2	-11	1	-17	22	0	28	4	-13	-25	3	6	-9	-3
2 PREVYR	29	100	-26	-27	0	7	-2	3	2	6	9	9	13	18	13	-10	-7	-21	7	-17	2	10
3 PROGGRP	-39	-21	100	93	-32	30	8	21	25	29	-5	26	-18	-3	-2	3	26	34	-3	-37	25	-14
4 PROGYX	-39	-26	93	100	-35	2[illegible]	-5	19	22	32	1	23	-19	2	-7	2	33	50	-4	-31	26	-11
5 R-ETT	4	14	-20	-22	100	-2	8	-20	-20	-11	-22	-14	-17	-13	-15	5	0	-18	-8	21	-46	7
6 R-DGI	4	-11	28	23	-18	100	-8	46	55	15	10	10	4	-19	17	2	1	-2	-4	-52	4	18
7 R-QFL	6	-13	-4	-10	35	34	100	-6	4	-26	-37	-10	7	-26	20	21	0	-23	0	21	6	-42
8 R-ADC	-3	-7	0	3	-1	-4	14	100	34	36	10	1	17	-9	29	0	0	10	-8	-33	34	4
9 R-FFD	12	20	20	14	-27	37	-11	-1	100	22	-4	5	4	-21	33	-5	0	1	4	-51	17	1
10 R-SMT	-3	12	16	9	-31	-19	-36	-25	13	100	5	-3	-3	-12	24	-2	0	15	6	-42	25	-15
11 R-NST	10	-4	-33	-26	11	-24	18	0	-48	-10	100	8	15	38	-4	-21	-1	20	0	-5	16	63
12 C-ACM	26	28	-17	-15	1	-20	-7	16	27	-2	-12	100	2	52	-41	-61	35	0	0	-27	8	48
13 C-QFL	21	20	-23	-22	19	-26	0	14	-2	2	19	23	100	0	32	2	-6	0	-3	4	21	4
14 C-STW	6	3	-34	-30	17	-25	0	-3	-30	2	19	16	-1	100	-34	-47	24	-2	3	-2	17	54
15 C-FMT	8	12	26	15	-14	34	8	-2	42	17	-34	-10	15	-37	100	6	-23	0	18	-20	26	-41
16 C-GRV	-10	2	-3	1	14	-2	2	-8	-14	-2	31	-52	6	3	-12	100	-17	2	-7	20	-10	-48
17 R-MISS	-28	-21	31	41	-1	0	0	0	0	0	0	1	4	-7	3	-1	100	54	-13	2	20	10
18 C-MISS	-24	-22	30	42	-9	-6	-8	-1	-7	5	1	0	0	0	1	-1	68	100	-12	12	32	-1
19 PRESFRY	1	7	-15	-10	18	-16	5	8	-1	-7	13	8	8	8	-4	5	-18	-7	100	1	18	3
20 SITEC5	7	1	-34	-24	45	-21	18	1	-45	-51	41	0	23	16	-33	18	0	14	13	100	0	-3
21 SITE35	-8	29	24	26	1	-10	-32	-15	12	35	-20	7	15	11	10	9	28	32	-5	4	100	0
22 SITE12	-3	-17	-13	-14	16	-24	5	-10	-30	14	28	0	2	53	-32	0	-10	-5	18	-2	0	100

Table 14

Correlation Matrix Showing Extent and Pattern of Collinearities
Among Predictors for SENTENCE READING (IRA3 - E)
Year 1 above diagonal; Year 2 below diagonal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 ENGCATG	100	34	-41	-43	10	-6	16	-7	-9	-16	0	1	14	15	21	-5	-2	-25	-2	6	-12	1
2 PREVYR	32	100	-29	-29	0	-1	8	-10	0	2	0	10	5	14	22	-6	-12	-23	12	-7	0	-2
3 PROGGRP	-38	-18	100	94	-44	23	-18	31	21	37	9	3	3	-2	6	6	25	54	0	-33	37	-20
4 PROGYx	-38	-25	93	100	-41	22	-26	29	23	39	10	2	-4	-2	1	7	31	64	-2	-3	32	-17
5 R-ETT	7	0	-28	-28	100	-4	3	-28	-24	-11	-22	-17	-21	-14	-28	3	-2	-20	-10	23	-48	16
6 R-DGI	1	-2	36	29	-16	100	-32	39	47	17	26	18	14	-10	7	-4	10	20	[illegible]	-57	5	32
7 R-QFL	9	-6	-12	-16	33	34	100	-20	-14	-37	-29	-7	15	-14	9	14	1	-27	-5	33	6	-44
8 R-ADC	-4	-16	-8	-2	7	-13	14	100	31	30	30	17	31	2	23	-6	6	31	-14	-27	40	20
9 R-FRD	11	21	16	12	-30	30	-21	-11	100	19	6	15	12	-1	27	-15	11	18	-4	-45	24	14
10 R-SMT	-2	16	22	13	-36	-24	-39	-28	15	100	19	-1	-4	-6	20	-5	7	22	5	-3	27	-4
11 R-NST	12	11	-31	-26	10	-16	26	10	-39	-12	100	4	9	38	30	-20	-5	18	4	-26	20	58
12 C-ACM	25	30	-18	-15	-2	-21	-9	18	32	-2	-14	100	42	11	14	-54	10	15	-4	-18	17	23
13 C-QFL	22	15	-33	-29	23	-30	1	13	-1	2	26	23	100	9	38	-7	1	-4	-3	-8	25	4
14 C-STW	6	11	-26	-27	20	-20	6	9	-17	1	8	18	6	100	20	-29	1	-2	10	-5	24	31
15 C-PMT	10	7	11	5	-12	25	-2	-26	16	22	-14	-14	10	-14	100	-33	5	5	21	-22	37	[illegible]
16 C-GRV	-11	2	5	6	16	-2	5	5	-12	-1	31	-54	8	-7	-2	100	-3	-3	-8	17	-11	-38
17 R-MISS	-29	-19	36	44	-3	4	1	2	7	1	-5	1	5	-14	14	-1	100	64	-8	4	24	-14
18 C-MISS	-25	-23	34	45	-13	-3	-8	1	-3	7	-3	0	0	-4	10	-1	68	100	-7	-1	36	4
19 PRESFRY	-1	6	-18	-11	19	-20	4	8	-7	-8	18	8	7	12	-11	5	-17	-7	100	4	20	4
20 SITE05	9	-11	-36	-26	48	-12	25	11	-39	-54	35	0	29	7	-17	15	-5	11	16	100	-1	-16
21 SITE35	-8	6	28	29	-3	-10	-35	-17	16	37	-22	8	13	8	15	9	28	32	-5	1	100	0
22 SITE12	-2	5	-10	-14	12	-16	9	-3	-19	17	20	-1	5	48	-17	-3	-14	-8	21	-14	0	100

Table 15

Correlation Matrix Showing Extent and Pattern of Collinearities
Among Predictors for NARRATIVE READING COMPREHENSION (IRAS - E)
Year 1 above diagonal; Year 2 below diagonal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 ENGCATG	100	32	-38	-43	9	1	13	1	-2	-11	0	-1	22	-1	28	4	-13	-25	3	6	-10	-4
2 PREVYR	38	100	-27	-23	0	7	-2	2	2	6	7	8	13	17	13	-9	-8	-22	8	-17	1	9
3 PROGGRP	-37	-26	100	94	-32	30	8	21	24	29	-5	26	-18	-2	-2	3	26	34	-3	-37	25	-14
4 PROGYx	-38	-29	93	100	-35	24	-5	19	21	33	1	24	-19	2	-7	2	34	51	-4	-30	26	-10
5 R-ETT	4	3	-21	-23	100	-2	8	-20	-20	-11	-22	-14	-17	-13	-15	5	0	-18	-8	21	-46	7
6 R-DGI	4	-1	29	24	-18	100	-9	56	16	10	10	5	-19	18	1	1	1	-2	-5	-52	4	17
7 R-QFL	6	1	-5	-11	36	33	100	-6	3	-26	-36	-10	6	-25	19	21	1	-23	0	21	6	-42
8 R-ADC	-3	-3	-1	2	0	-4	15	100	34	30	11	1	17	[illegible]	29	-1	1	11	-8	-33	34	4
9 R-FRD	11	27	22	15	-28	37	-13	-11	100	22	-4	5	5	-22	33	-6	0	1	4	-51	18	0
10 R-SHT	-2	8	17	11	-33	-18	-37	-26	15	100	6	-3	-3	-12	24	-3	0	15	6	-42	25	-15
11 R-HST	9	1	-34	-27	11	-25	18	1	-49	-10	100	8	15	38	-4	-21	-1	20	-1	-5	16	63
12 C-ACM	24	41	-16	-14	1	-20	-7	16	27	-2	-13	100	3	52	-41	-61	35	0	-1	-27	9	48
13 C-QFL	21	22	-25	-24	20	-27	2	14	-4	0	20	23	100	0	33	2	-6	0	-3	3	21	4
14 C-STW	6	8	-35	-32	19	-25	2	-2	-31	-1	19	16	1	100	-35	-47	25	-1	2	-2	17	54
15 C-FMT	8	11	26	15	-14	34	8	-2	42	16	-34	-10	14	-36	100	6	-23	-1	19	-20	26	-41
16 C-GRV	-8	-8	-4	0	13	-3	-2	-8	-14	-1	32	-53	7	1	-11	100	-17	2	-7	20	-10	-48
17 R-MISS	-27	-18	32	42	-2	0	0	0	1	0	0	1	4	-9	3	0	100	54	-13	2	21	10
18 C-MISS	-23	-19	31	44	-10	-6	-9	-2	-7	6	1	0	0	0	1	0	67	100	-13	12	32	-1
19 PRESFRY	0	11	-15	-10	17	-15	5	8	-1	-7	13	8	8	9	-4	4	-18	-7	100	1	18	3
20 SITE05	6	-9	-36	-26	46	-22	19	2	-46	-52	42	0	25	18	-33	17	1	15	13	100	-1	-2
21 SITE35	-8	3	25	27	-2	-9	-33	-16	13	36	-20	8	13	7	11	9	29	33	-5	1	100	0
22 SITE12	-2	0	-13	-14	15	-25	5	-10	-30	14	29	-1	3	53	-32	0	-13	-7	18	-1	0	100

99

Table 16

Correlation Matrix Showing Extent and Pattern of Collinearities
Among Predictors for EXPOSITORY READING COMPREHENSION (IRAS - E)
Year 1 above diagonal; Year 2 below diagonal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 ENGCATG	100	35	-42	-43	11	-7	16	-7	-9	-17	1	0	14	16	20	-4	-2	-26	-1	7	-11	2
2 PREVYR	27	100	-29	-29	3	0	7	-10	-1	1	1	8	5	13	21	-4	-11	-22	12	-6	-1	0
3 PROGGRP	-38	-16	100	94	-45	23	-18	31	21	37	9	4	3	-1	6	6	25	54	0	-33	38	-20
4 PROGYx	-38	-19	93	100	-43	22	-26	29	23	40	10	3	-4	-2	2	7	3	64	-2	-29	33	-18
5 R-ETT	7	4	-28	-28	100	-5	5	-29	-23	-11	-25	-15	-21	-14	-26	1	-2	-21	-8	22	-48	15
6 R-DGI	1	1	36	29	-16	100	-32	39	47	17	26	19	14	-10	7	-4	9	20	-13	-57	6	32
7 R-QFL	9	4	-12	-16	33	34	100	-20	-14	-38	-29	-0	15	-14	7	15	1	-27	-4	35	6	-44
8 R-ADC	-4	-16	-8	-2	7	-13	14	100	31	30	30	18	31	2	24	-7	6	31	-14	-28	40	20
9 R-PRD	11	10	16	12	-30	30	-21	-11	100	19	7	14	12	-2	27	-15	12	18	-5	-45	23	15
10 R-SMT	-2	3	22	13	-36	-24	-39	-28	15	100	20	-2	-4	-6	20	-4	7	23	6	-34	27	-3
11 R-NST	12	9	-31	-26	10	-16	26	10	-39	-12	100	7	9	38	33	-22	-5	18	4	-29	21	58
12 C-ACM	25	23	-18	-15	-2	-21	-9	18	32	-2	4	100	42	11	11	-54	11	17	-5	-17	17	25
13 C-QFL	22	11	-33	30	23	-30	1	13	-10	2	[illegible]	22	100	9	38	-7	1	-4	-3	-8	25	5
14 C-STW	5	8	-26	-27	20	-20	8	9	-17	1	8	18	6	100	20	-29	1	-2	10	-5	24	32
15 C-PMT	10	-1	11	5	-12	25	-2	-26	16	22	-14	-14	10	-14	100	-32	5	6	21	-21	38	4
16 C-GRV	-11	3	5	6	16	-2	5	-5	-12	-1	31	-54	8	-7	-2	100	-3	-3	-8	16	-12	-40
17 R-MISS	-29	-11	36	4	-3	4	1	2	7	1	-5	1	5	-14	14	-1	100	64	-8	4	25	-15
18 C-MISS	-24	-14	34	45	-13	-3	-8	1	-3	7	-3	0	0	-4	10	-1	68	100	-7	-1	36	4
19 PRESPRY	-1	6	-18	-11	19	-20	4	8	-7	-8	18	8	7	12	-11	5	-17	-7	100	5	20	4
20 SITE05	9	3	-36	-26	48	-12	25	11	-39	-54	35	[illegible]	29	7	-17	15	-5	11	16	100	-1	-18
21 SITE35	-9	3	28	29	-3	-10	-35	-17	16	37	-22	8	13	8	15	9	28	32	-5	1	100	0
22 SITE12	-2	-2	-10	-14	12	-16	9	-3	-19	17	20	-1	5	48	-17	-3	-14	-8	21	-144	0	100

Table 17

Correlation Matrix Showing Extent and Pattern of Collinearities
Among Predictors for NARRATIVE LISTENING COMPREHENSION <IRAS - E>
Year 1 above diagonal; Year 2 below diagonal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 ENGCATG	100	32	-38	-43	9	1	13	2	-2	-10	0	-17	22	-1	29	4	-13	-25	4	6	-9	-4
2 PREVYR	46	100	-27	-28	1	7	-2	2	2	5	7	8	13	17	13	-9	-8	-22	7	-17	1	9
3 PROGGRP	-37	-38	100	94	-31	30	9	21	25	29	-5	26	-19	-2	-2	-2	26	34	-4	-37	25	-14
4 PROGYx	-37	-35	93	100	-35	24	-5	19	21	32	1	23	-19	2	-8	2	34	50	-5	-30	26	-10
5 R-ETT	4	11	-21	-23	100	-2	8	-20	-20	-10	-22	-13	-16	-13	-15	5	0	-18	-7	21	-45	7
6 R-DGI	4	-26	29	24	-18	100	-9	47	56	16	10	10	5	-19	18	1	1	-1	-5	-53	5	17
7 R-OFL	6	-14	-5	-11	36	33	100	-6	3	-26	36	-10	7	-25	20	22	1	-23	0	21	6	42
8 R-ADC	-3	-4	-1	2	0	-4	15	100	34	30	11	1	17	-9	29	-1	1	11	-9	-33	34	4
9 R-PFD	11	0	22	15	-28	37	-13	-2	100	23	-4	6	5	-22	34	-5	1	1	4	-52	18	0
10 R-SMT	-2	-11	17	11	-33	-18	-37	-26	15	100	5	-3	-3	-12	24	-3	0	15	5	-42	25	-15
11 R-NST	10	14	-34	-28	11	-25	18	1	-49	-11	100	8	14	38	-5	-21	-1	20	-1	-4	16	63
12 C-ACM	24	20	-16	-14	1	-20	-7	16	27	-2	-13	100	2	52	-41	-61	35	0	1	-27	8	48
13 C-OFL	21	25	-25	-24	21	-27	2	15	-4	-1	20	23	100	0	32	1	-6	0	4	4	21	4
14 C-STW	5	5	-35	-32	19	-25	2	-2	-31	-1	19	16	2	100	-35	-47	25	-1	2	-2	17	54
15 C-PMT	8	-9	25	15	-14	34	8	-2	42	16	-34	-10	14	-36	100	6	-23	-1	18	-20	25	-41
16 C-GRV	-8	-4	-4	0	13	-3	2	-8	-14	-1	32	-53	7	1	-11	100	-17	2	-7	20	-10	-48
17 R-MISS	-28	-16	31	41	-1	0	0	0	1	0	1	1	4	-9	3	0	100	54	-13	2	21	10
18 C-MISS	-24	-19	29	42	-10	-6	-9	-2	-7	6	1	0	-1	0	1	0	68	100	-13	12	32	-1
19 PRESFRY	1	-2	-16	-10	18	-15	5	8	-1	-7	13	9	7	9	-4	4	-18	-7	100	1	17	3
20 SITE05	6	23	-35	-26	46	-21	19	2	-46	-52	42	0	25	18	-33	17	-1	14	13	100	0	-2
21 SITE35	-7	-8	25	27	-2	-9	-33	-16	14	36	-21	8	12	7	11	9	28	32	-5	2	100	0
22 SITE12	-3	-11	-13	-15	15	-24	5	-10	-30	14	29	-1	3	52	-31	0	-10	-5	18	-2	0	100

Table 18

Correlation Matrix Showing Extent and Pattern of Collinearities
Among Predictors for EXPOSITORY LISTENING COMPREHENSION (IRAS - E)
Year 1 above diagonal; Year 2 below diagonal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 ENGCATG	100	35	-41	-43	10	-6	17	-7	-9	-15	0	1	14	15	21	-5	-2	-26	-1	6	-11	1
2 PREVYR	44	100	-30	-30	1	-1	10	-10	1	3	-1	10	5	14	23	-6	-12	-23	13	-7	1	-2
3 PROGGRP	-38	-29	100	94	-44	23	-18	31	20	36	9	3	3	-2	5	6	25	54	0	-33	36	-20
4 PROGYx	-38	-31	93	100	-41	22	-26	29	23	39	10	2	-4	-2	1	7	31	64	-2	-28	32	-17
5 R-ETT	7	5	-28	-28	100	-4	4	-28	-24	-11	-22	-17	-21	-14	-28	3	-2	-20	-9	23	-47	16
6 R-DGI	1	-7	36	29	-16	100	-31	39	47	17	26	18	14	-10	7	-4	10	20	-13	-57	5	32
7 R-QFL	9	-3	-12	-16	33	34	100	-20	-13	-36	-30	-7	15	-13	9	13	0	-28	-4	33	7	-44
8 R-ADC	-4	-11	-8	-2	7	-13	14	100	31	30	30	17	31	2	23	-6	6	31	-14	-27	40	20
9 R-PRD	11	9	16	12	-30	30	-21	-11	100	19	6	15	12	-1	27	-15	11	17	-4	-45	24	14
10 R-SMT	-2	1	22	13	-36	-24	-40	-28	15	100	19	-1	-4	-6	21	-5	7	22	6	-34	27	-4
11 R-NST	12	12	-31	-26	10	-16	26	10	-39	-12	100	4	9	38	30	-20	-5	18	4	-26	20	58
12 C-ACM	25	9	-18	-15	-2	-21	-9	18	32	-2	-14	100	42	11	14	-54	10	15	-4	-18	16	23
13 C-QFL	22	15	-33	-29	23	-30	1	13	-10	2	26	23	100	9	38	-7	1	-4	-3	-8	25	4
14 C-STW	5	7	-26	-27	20	-20	8	9	-17	1	1	18	6	100	20	-29	1	-2	10	-5	24	31
15 C-PMT	10	2	11	5	-12	25	-2	-26	16	22	-14	-14	10	-14	100	-33	5	5	21	-22	37	2
16 C-GRV	-11	2	5	6	16	2	5	-5	-12	-1	31	-54	84	-7	-2	100	-2	-3	-8	17	-12	-38
17 R-MISS	-29	-22	36	44	-3	4	1	24	7	1	-5	1	5	-14	14	-1	100	64	-8	4	24	-14
18 C-MISS	-24	-30	34	45	-13	-3	-8	1	-3	7	-3	0	0	-4	10	-1	68	100	-7	-1	35	4
19 PRESPRY	-1	-6	-18	-11	19	-20	4	8	-7	-8	18	8	7	12	-11	5	-17	-7	100	4	20	4
20 SITE05	9	-1	-36	-26	48	-12	25	11	-39	-54	35	0	29	7	-17	15	-5	11	16	100	0	-16
21 SITE35	-8	-4	28	29	-3	-10	-35	-17	16	37	-22	8	13	8	15	9	28	32	-5	1	100	0
22 SITE12	-2	-5	-10	-14	12	-16	9	-3	-19	17	20	-1	5	48	-17	-3	-14	-8	21	-14	0	100

less than .40. Correlations with single word decoding and Sentence Reading drop to about .30.

In general, previous achievement is related to program assignment (number of years in Spanish reading instruction) at a level of about -.3 over all IRAS measures and for both Years 1 and 2. The pattern is one in which students whose performance on IRAS-E is below par in the previous year are more likely to be kept in Spanish reading instruction programs. The pattern is virtually without exception in Year 1, but there are some intriguing departures during Year 2. In particular, the strength of this relation becomes more noticeable for the oral IRAS measures (Vocabulary Definition at -.45; Narrative Listening Comprehension at -.38), and weaker for the direct measures of decoding (Letter-sound Decoding at -.17; Letter-sound Spelling at -.21; Sentence Reading at -.18). The changes are not huge, but the trend is for assignment to Spanish reading instruction to depend more in Year 2 on spoken language and less on decoding level.

The other differences between tables can be seen in comparing Sentence Reading (Table 14) and Narrative Reading Comprehension (Table 15). Sentence Reading was added to the IRAS battery with the third cohort of students, whereas Narrative Reading was in the original battery. The result is that most students from the first cohort are not represented in Table 14 (those who entered at first grade), with consequent changes in the structure of the instructional variables (relatively slight), and the program and site contrasts (relatively larger, because the first cohort was exclusively from one of the border sites). While it is important to take note of these variations, none of them are of substantial magnitude.

Predictors for Spanish Reading

The collinearities for predictors of deviations in Spanish IRAS performance are shown in Tables 19 to 27, following the same pattern as for English IRAS. As noted earlier in the volume, the limited amount of Spanish instruction meant that many of the instructional variables required estimation, the consequence being an artifactual reduction in the predictive power of these variables, but also in the extent of collinearity.

Precursor predictors. Recall that the entry language category for both Spanish and English are based on teacher ratings of the student's oral language fluency on entry to kindergarten. The correlation between this meaure (in Spanish) and the index of previous achievement standing on entry to Year 1 (Alphabet Knowledge on entry to Year 1) is about .20. Students typically did not acquire knowledge of the alphabet as a precursor to Spanish reading, and so the weakness of this relation was not unexpected.

The relation of entry language category to previous achievement in Year 2 is more interesting. As was true for English, the highest correlations are with the oral language components of IRAS; Vocabulary Definition, and both Narrative and Expository Listening Comprehension

Table 19

Correlation Matrix Showing Extent and Pattern of Collinearities
Among Predictors for VOCABULARY DECODING (IRAS - S)
Year 1 above diagonal; Year 2 below diagonal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 SPHCATG	100	22	31	29	-4	16	8	-3	7	13	-1	-2	-5	5	-4	9	-20	-24	13	-37	41	4
2 PREVYR	31	100	21	12	-19	27	-5	-26	6	6	2	4	-9	16	3	-10	-20	-15	2	-26	12	5
3 FROGGRP	32	14	100	94	2	18	-11	-5	12	-6	10	1	-7	9	-1	-3	-60	-74	-3	-36	25	-14
4 PROGYx	29	16	93	100	8	13	-6	10	15	-9	4	-1	4	1	2	0	-68	-85	-4	-30	26	-10
5 R-GFL	-3	-4	-1	-2	100	-10	-37	-10	34	-36	28	-11	29	-8	-24	-29	-1	-9	8	26	22	-25
6 R-DGI	-3	-21	12	6	26	100	-16	10	44	-25	-16	9	-9	23	1	46	1	-13	4	-33	-16	20
7 R-ETT	-5	11	-10	-2	38	-17	100	-14	-34	28	-25	-3	12	4	36	8	-2	6	-6	4	-9	25
8 R-NST	5	-8	-1	7	-37	-29	-45	100	32	6	-39	19	15	-9	5	65	-1	-11	-6	-6	1	17
9 R-HDC	-5	-4	-5	-4	30	17	45	-49	100	-26	-1	25	36	28	-15	24	-1	-15	17	-8	19	8
10 R-SHT	8	7	-3	0	-7	-3	18	16	17	100	-12	-8	2	-1	22	2	-2	9	-1	-6	-1	2
11 R-CNT	2	6	-5	1	-8	-46	23	36	-4	33	100	-18	-18	-21	-18	-25	0	-4	2	-4	-6	-36
12 C-ADC	-3	-9	8	-3	18	29	-3	-43	15	-8	-19	100	14	43	-11	7	-4	0	-3	-6	8	28
13 C-STW	3	6	4	5	14	-2	-4	3	-11	-3	2	-9	100	10	4	-10	-19	-2	10	27	25	16
14 C-FMT	5	21	0	-3	16	14	24	-29	13	-10	-10	10	1	100	16	-1	10	0	18	-9	17	24
15 C-DTC	6	6	0	0	-11	12	-20	22	-6	17	3	-31	35	-12	100	8	0	-1	[illegible]	-6	-6	16
16 C-NST	9	23	1	4	[illegible]	-34	16	20	-9	32	52	-19	23	-11	16	100	4	0	-3	-25	-9	18
17 R-MISS	-23	-16	-68	-75	1	-8	12	-6	[illegible]	-1	1	15	-26	11	-31	13	100	81	1	2	-33	-12
18 C-MISS	-30	-22	-86	-93	1	-7	2	-5	4	-1	1	2	1	1	-2	0	76	100	1	14	-28	-5
19 PRESPRY	-9	6	-15	-9	-12	-1	-6	6	-2	3	2	-9	17	19	4	-1	-3	10	100	1	18	3
20 SITE05	-39	-8	-35	-28	16	-12	29	-2	11	3	12	-20	22	5	-1	15	14	28	11	100	0	-2
21 SITE35	43	27	26	28	3	-2	13	1	13	7	8	-15	7	4	8	13	-22	-28	-7	0	100	0
22 SITE12	3	-22	-13	-15	1	-15	-25	32	-15	-14	2	-25	24	-16	22	4	3	22	18	-2	0	100

899

Table 20

Correlation Matrix Showing Extent and Pattern of Collinearities
Among Predictors for VOCABULARY DEFINITION <IRAS - 5>
Year 1 above diagonal; Year 2 below diagonal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 SPHCATG	100	18	31	29	4	11	3	-2	7	11	5	-9	-5	-2	-9	8	-18	-22	13	-34	43	-9
2 PREVYR	44	100	11	2	-1	21	-23	-22	9	-2	19	-13	-10	-1	-10	-7	-10	-4	3	-18	14	-21
3 PROGGRP	32	35	100	93	20	13	-24	-1	13	-13	23	-11	-7	-4	-11	-1	-56	-72	-3	-35	27	-30
4 PROGYx	28	31	93	100	26	5	-17	30	15	-15	16	-14	4	-12	-7	-1	-65	-83	-4	-28	28	-25
5 R-QFL	-3	5	-1	-2	100	4	-25	-17	41	-31	13	10	34	15	-12	-36	-20	-30	9	16	25	-3
6 R-DGI	-3	-7	12	6	26	100	-28	9	46	-31	-6	-10	-11	16	-7	45	13	-4	3	-29	-17	8
7 R-ETT	-5	2	-10	-2	38	-17	100	-10	-37	24	-15	-19	12	-12	28	12	11	19	-6	15	-10	12
8 R-NST	5	-5	-1	7	37	29	-45	100	30	10	-43	26	15	-4	10	62	-4	-14	-7	-9	1	26
9 R-ADC	-5	-8	-5	-4	30	17	45	-49	100	-26	-15	27	36	33	-15	22	0	-16	18	-8	19	10
10 R-SMT	8	0	-3	0	-7	-3	18	16	17	100	-7	-17	2	-10	18	5	4	16	0	-2	-1	-8
11 R-CNT	2	4	-5	1	-8	-46	23	36	-4	33	100	-4	-20	-9	-8	-29	-14	-19	3	-15	-7	-25
12 C-ADC	-3	-3	8	-3	18	29	-3	-43	15	-8	-19	100	15	33	-26	9	10	14	-3	5	9	13
13 C-STW	3	-3	4	5	14	-2	-4	3	-11	-3	2	-9	100	11	5	-13	-21	-1	10	28	25	20
14 C-PMT	5	23	0	-3	16	14	24	-29	13	-10	-10	10	1	100	4	2	27	15	19	1	19	7
15 C-DTC	6	-6	0	0	-11	12	-20	22	-6	17	3	-31	35	-12	100	12	11	10	-5	2	-6	6
16 C-NST	10	6	0	4	-4	-34	16	20	-9	32	52	-19	23	-11	16	100	5	1	-4	-28	-10	23
17 R-MISS	-23	-23	-68	-75	1	-8	2	-6	5	-1	1	15	-26	11	31	-13	100	78	1	-5	-35	-1
18 C-MISS	-30	-41	-86	-93	1	-7	2	-5	4	-1	1	2	1	1	-2	0	76	100	1	9	-30	9
19 PRESPRY	-9	-8	-15	-9	-12	-1	-6	6	-2	3	2	-9	17	19	4	-1	-3	10	100	2	18	1
20 SITE05	-39	-24	-35	-26	16	-12	29	-2	11	0	12	-20	22	5	-1	15	13	26	11	100	0	24
21 SITE35	42	30	25	28	3	-2	13	6	13	8	8	-15	7	3	8	14	-22	-28	-2	1	100	0
22 SITE12	3	-31	-14	-15	1	-15	-25	32	-15	-14	2	-25	24	-16	22	4	3	22	18	-2	0	100

Table 21

Correlatio Matrix Showing Extent and Pattern of Collinearities
Among Predictors for LETTER-SOUND DECODING (IRAS - 5)
Year 1 above diagonal; Year 2 below diagonal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 SPNCATG	100	25	31	31	-4	16	8	-3	7	13	-1	-2	-5	5	-4	9	-21	-25	12	-40	41	4
2 PREVYR	35	100	22	14	-20	27	-5	-27	6	6	2	5	-9	16	32	-11	-22	-17	2	-27	12	6
3 PROGGRP	32	14	100	94	2	18	-11	-5	12	-6	10	1	-7	9	-1	3	-60	-74	-3	-37	27	-14
4 PROGYx	29	16	93	100	8	13	-6	10	15	-9	4	-2	4	1	2	0	-68	-85	-4	-30	30	-11
5 R-OFL	-3	-6	-1	-2	100	-10	-37	-10	34	-36	28	-11	29	-8	-24	-29	-1	-9	8	26	23	-25
6 R-DGI	-3	-24	12	6	26	100	-16	10	44	-25	-16	9	-9	23	1	46	1	-13	4	-34	-16	20
7 R-ETT	-5	14	-10	-2	38	-17	100	-14	-34	28	-25	-3	12	4	36	8	-2	6	-6	4	-9	25
8 R-NST	5	-7	-1	7	-37	-29	-45	100	32	6	-35	19	15	-9	5	65	-1	-11	-6	-6	1	17
9 R-ADC	5	-8	-5	-4	30	17	45	-49	100	-26	-14	25	36	26	-15	24	-1	-15	17	-8	19	8
10 R-SMT	8	5	-3	0	-7	-3	18	16	17	100	-12	-8	2	1	22	2	-2	9	-1	-6	-1	2
11 R-CNT	2	8	-5	1	-8	46	23	36	-4	33	100	-18	-18	-22	-18	-25	0	-4	2	-4	-6	-36
12 C-ADC	-3	-18	8	-3	18	29	-3	-43	15	-8	-19	100	14	43	-11	7	-4	0	-3	-6	8	28
13 C-STW	3	8	4	5	14	-2	-4	3	-11	-3	2	-9	100	10	4	-10	-19	-2	10	27	25	16
14 C-FMT	5	23	0	-3	16	14	24	-29	13	-10	-10	10	1	100	16	-1	10	0	18	-9	17	24
15 C-DTC	6	6	0	0	-11	12	-20	22	-6	17	3	-31	35	-12	100	8	0	-1	-5	-6	-6	18
16 C-NST	9	23	1	4	-4	-34	16	20	-9	32	52	-19	23	-11	16	100	4	0	-3	-25	-10	18
17 R-MISS	-23	-19	-68	-75	1	-8	2	-6	5	-1	1	15	-26	11	-31	-13	100	81	1	1	-36	-12
18 C-MISS	-30	-22	-86	-93	1	-7	2	-5	4	-1	1	2	1	1	-2	0	76	100	0	14	-31	-4
19 PRESFRY	-10	2	-15	-9	-12	-1	-6	6	-2	3	2	-9	17	19	4	-1	-3	10	100	0	16	3
20 SITE05	-39	-5	-36	-26	16	-12	29	-2	11	0	12	-20	22	5	1	15	14	28	12	100	-2	-3
21 SITE35	43	29	26	28	3	-2	13	1	13	7	8	-15	7	4	8	13	-23	-28	-3	0	100	1
22 SITE12	3	-24	-13	-15	1	-15	-25	32	-15	-14	2	-25	24	-16	22	4	3	22	18	-2	0	100

Table 22

Correlation Matrix Showing Extent and Pattern of Collinearities
Among Predictors for LETTER-SOUND SPELLING <IRAS - 5>
Year 1 above diagonal; Year 2 below diagonal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 SPNCATG	100	22	31	30	-4	16	8	-3	7	13	-1	-2	-5	5	-4	9	-21	-24	13	-38	41	4
2 PREVYR	32	100	21	12	-19	-7	-5	-26	6	6	2	4	-9	16	3	-10	-20	-15	2	-26	13	5
3 PROGGRP	33	17	100	94	2	18	-11	-5	12	-6	10	1	-7	9	-1	-3	-60	-75	-3	-36	26	-14
4 PROGYx	19	15	93	100	8	12	-6	10	14	-9	4	-2	4	1	2	0	-68	-85	-5	-30	27	-10
5 R-QFL	-2	-3	-2	-3	100	-10	-37	-10	34	-36	28	-11	29	-6	-24	-29	-1	-9	8	26	22	-25
6 R-DGI	-2	-13	11	5	25	100	-16	10	44	-25	-16	9	-9	23	1	46	1	-13	4	-33	-16	20
7 R-ETT	-5	11	-10	-2	39	-17	100	1	-34	28	-25	-3	12	4	36	8	-2	6	-6	4	-9	25
8 R-NST	4	-12	1	8	-36	-28	-46	100	32	6	35	19	15	-9	5	65	-1	-11	-6	-6	1	17
9 R-ADC	-5	-5	-5	-4	31	17	45	-50	100	-26	-14	25	36	28	-15	24	-1	-15	17	-8	19	8
10 R-SMT	8	10	-3	0	-7	-3	18	16	17	100	-12	-8	2	-1	22	2	-2	9	-1	-6	-1	2
11 R-CNT	1	5	-4	2	-7	-45	23	34	-4	33	100	-16	-18	-22	-18	-25	0	-4	2	-4	-6	-36
12 C-ADC	-2	-7	7	-5	16	26	-3	-42	15	-8	-17	100	14	43	-11	7	-4	0	-3	-6	8	28
13 C-STW	3	16	4	5	14	-2	-4	3	-11	-3	2	-10	100	10	4	-10	-19	-2	10	27	25	16
14 C-PMT	4	19	1	-2	18	16	24	-32	13	-10	-11	12	1	100	16	-1	10	0	18	-9	17	24
15 C-DTC	6	7	0	0	-12	12	-20	23	-6	17	4	-32	35	-11	100	8	0	-1	-5	-6	-6	18
16 C-NST	10	22	0	4	-5	-35	16	21	-9	32	53	-21	23	-10	16	100	4	0	-3	-25	-5	18
17 R-MISS	-24	-17	-68	-75	2	-7	2	-8	5	-1	0	17	-26	10	-31	-12	100	81	1	2	-33	-12
18 C-MISS	-31	-18	-86	-93	2	-7	2	-7	4	-1	0	4	1	0	2	1	76	100	1	15	-28	-5
19 PRESFRY	-10	6	-14	-8	-11	-1	-6	5	-2	3	1	-8	17	18	4	-1	-4	9	100	1	17	3
20 SITE05	-39	-7	-34	-25	17	-12	29	-3	11	0	11	-19	22	4	-1	15	12	27	11	100	0	-2
21 SITE35	43	31	25	27	4	-2	13	1	13	8	8	-15	7	4	8	14	-22	-27	-2	2	100	0
22 SITE12	2	-22	-13	-14	2	-15	-25	31	-15	-14	1	-24	25	-17	22	4	2	22	17	-2	0	100

Table 23

Correlation Matrix Showing Extent and Pattern of Collinearities
Among Predictors for SENTENCE READING (IRAS - S)
Year 1 above diagonal; Year 2 below diagonal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 SPNCATG	100	17	35	32	-2	5	17	9	8	19	-8	-7	-1	-5	-4	13	-22	-26	11	-34	48	7
2 PREVYR	32	100	-5	-9	-7	0	4	-7	7	12	-7	-11	3	-2	0	0	3	10	7	2	19	-6
3 PROGGRP	34	25	100	94	29	-4	-1	22	16	-13	-3	-12	15	-17	-6	5	-61	-77	0	-32	37	-20
4 PROGYx	30	24	93	100	32	-3	-5	31	19	-19	-3	-12	19	-21	-4	6	-68	-87	-3	-28	32	-17
5 R-OFL	-6	-9	-6	-6	100	-7	-50	-3	42	-23	-3	16	39	14	-8	-41	-22	-35	10	21	29	-10
6 R-DGI	-11	-26	-6	-7	22	100	1	47	50	-21	-44	1	5	17	10	65	24	4	1	-29	-18	29
7 R-ETT	-5	4	-10	-1	39	-18	100	-44	-40	26	21	-30	-10	-18	13	3	-9	4	-2	1	-14	-1
8 R-NST	12	-5	19	21	-36	-10	-52	100	45	-23	-26	29	6	-1	-2	70	-19	-34	-7	-31	1	33
9 R-ADC	-7	-9	-6	-4	28	19	45	-56	100	-31	-35	29	41	35	-9	29	-1	-20	18	-12	20	-13
10 R-SMT	8	-2	1	3	-9	2	17	12	15	100	10	-19	-11	-14	15	0	4	20	-2	-4	-1	-20
11 R-CNT	6	2	2	6	-4	-43	23	32	-2	34	100	9	-17	-19	10	-27	11	3	-4	5	-9	17
12 C-ADC	-7	-6	-4	-12	17	17	-2	-34	17	-4	-15	100	21	33	-30	8	2	11	-2	3	9	13
13 C-STW	2	15	0	3	13	-8	-4	9	-12	-2	4	-13	100	20	2	-14	-34	-17	11	23	28	13
14 C-PMT	0	27	-15	-13	12	-3	27	-16	14	-7	-3	-1	-4	100	4	1	29	24	23	13	20	19
15 C-DTC	5	9	-2	-1	-13	11	-20	26	-6	17	4	-34	35	-14	100	8	7	7	-2	-8	-7	6
16 C-NST	11	16	7	9	-1	-31	16	16	-8	32	52	-16	25	-6	18	100	-5	-7	-1	-44	-10	32
17 R-MISS	-23	-17	-64	-73	6	9	1	-24	5	-5	-5	28	24	25	-32	-19	100	77	3	-5	-39	-5
18 C-MISS	-32	-24	-86	-93	4	5	1	-18	4	-4	-4	11	3	11	-1	-3	74	100	1	10	-34	3
19 PRESPRY	-12	4	-17	-10	-15	-3	-7	8	-4	1	4	-10	17	19	4	-1	-3	10	100	4	20	4
20 SITE05	-36	-12	-36	-26	20	-2	30	-14	12	-3	9	-15	26	15	0	12	11	29	15	100	0	-16
21 SITE35	44	25	29	30	4	-3	13	1	13	8	8	-16	7	4	8	14	-24	-29	-3	0	100	0
22 SITE12	10	-11	-10	-13	4	-6	-27	26	-16	-18	-2	-21	28	-10	24	1	-1	23	21	-14	0	100

Table 24

Correlation Matrix Showing Extent and Pattern of Collinearities Among Predictors for NARRATIVE READING COMPREHENSION (IRAS - 5)
Year 1 above diagonal; Year 2 below diagonal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 SPHCATG	100	22	31	30	-4	16	8	-3	7	13	-1	-2	-5	5	-4	9	-21	-24	13	-38	41	4
2 PREVYR	24	100	21	12	-19	26	-5	-26	6	6	1	4	-9	16	3	-10	-20	-15	2	-26	13	4
3 PROGGRP	32	8	100	93	2	18	-10	-4	12	-6	10	1	-7	9	-1	-3	-60	-74	-3	-36	25	-14
4 PROGYx	29	8	93	100	8	12	-5	10	14	-9	4	-2	4	1	2	0	-68	-85	-4	-30	27	-10
5 R-GFL	-3	3	0	-2	100	-10	-37	-10	34	-35	28	-11	29	-7	-24	-29	-1	-9	7	26	22	-25
6 R-DGI	-3	3	12	6	25	100	-16	10	44	-25	-16	8	-9	23	1	46	1	-13	4	-33	-15	20
7 R-ETT	-5	3	-10	-1	38	-17	100	-14	-34	28	-25	-3	11	4	35	7	-2	6	-6	4	-9	25
8 R-HST	5	-8	-1	7	-37	-29	-45	100	32	6	-34	18	15	-9	5	65	-1	-10	-6	-6	1	17
9 R-ADC	-5	-1	-5	-4	30	17	44	-49	100	-26	-14	25	36	28	-15	24	0	-15	17	-9	19	8
10 R-SNT	8	-1	-3	0	-7	-3	18	15	17	100	-12	-8	2	-1	22	1	-2	9	-1	-6	-1	2
11 R-CNT	2	-2	-4	1	-8	-46	23	35	-4	33	100	-18	-18	-22	-17	-25	0	-4	2	-3	-6	-36
12 C-ADC	-3	0	8	-3	18	28	-3	·43	15	-8	-18	100	14	43	-11	7	-4	0	-3	-6	8	28
13 C-STW	3	5	4	5	14	-2	-4	3	-11	-3	2	-9	100	10	4	-10	-19	-1	10	26	25	16
14 C-PNT	5	20	0	-3	16	14	24	-29	13	-10	-10	10	0	100	16	-1	9	0	17	-9	17	24
15 C-DTC	6	0	0	0	-11	12	-20	22	-6	17	3	-31	35	-12	100	7	0	0	-5	-6	-6	18
16 C-HST	9	6	1	4	-4	-34	16	20	-9	32	52	-20	23	-11	16	100	4	0	-3	-25	-9	18
17 R-MISS	-23	-3	-68	-75	1	-8	2	-6	5	-1	1	15	-26	11	-31	-13	100	81	1	2	-33	-12
18 C-MISS	-30	-11	-86	-93	1	-7	2	-5	4	-1	1	2	0	1	-2	0	76	100	1	15	-28	-5
19 PRESFRY	-10	6	-15	-9	-12	-1	-6	6	-2	2	2	-9	17	19	4	-1	-2	10	100	1	17	3
20 SITE05	-39	-16	-36	-26	16	-12	29	-2	11	0	12	-20	-22	5	-1	14	14	28	11	100	0	-2
21 SITE35	43	11	26	28	3	-2	13	1	13	7	7	-15	7	3	8	13	-22	-28	-3	0	100	0
22 SITE12	3	-18	-13	-15	1	-15	-25	31	-15	-14	2	-25	24	-16	22	4	3	22	18	-1	0	100

Table 25

Correlation Matrix Showing Extent and Pattern of Collinearities
Among Predictors for EXPOSITORY READING COMPREHENSION (IRAS - 5)
Year 1 above diagonal; Year 2 below diagonal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 SPHCATG	100	17	35	32	-2	5	17	9	8	19	-8	-7	-1	-5	-4	13	-22	-26	11	-34	48	7
2 PREVYR	14	100	-5	-9	-7	0	4	-7	7	12	-7	-11	3	-2	0	0	3	10	7	2	19	-6
3 PROGGRP	34	15	100	94	29	-4	-1	22	16	-13	-3	-12	15	-17	-8	5	-61	-77	0	-32	37	-20
4 PROGYx	30	21	93	100	32	-3	-5	31	19	-19	-3	-12	19	-21	-4	6	-68	-87	-3	-28	32	-17
5 R-QFL	-6	-18	-6	-6	100	-7	-50	-3	42	-23	-3	16	39	14	-8	-41	-23	-35	10	21	29	-10
6 R-DGI	-11	-16	-6	-7	22	100	1	47	50	-21	-44	1	5	17	10	65	24	4	1	-29	-18	29
7 R-ETT	-5	-1	-10	-1	39	-18	100	-44	-40	26	21	-30	-10	-18	13	3	-9	4	-2	1	-14	-1
8 R-HST	12	-6	19	21	-36	-10	-52	100	45	-23	-26	29	6	-1	-2	70	-19	-34	-7	-31	1	33
9 R-ADC	-7	-9	-6	-4	28	19	45	-56	100	-31	-35	29	41	35	-9	29	-1	-20	16	-12	20	15
10 R-SMT	8	-13	1	3	-9	2	17	12	15	100	10	-19	-11	-14	15	0	4	20	-2	-4	-1	-13
11 R-CNT	6	1	2	6	-4	-43	23	32	-2	34	100	9	-17	-19	10	-27	11	3	-4	5	-9	-20
12 C-ADC	-7	-5	-4	-12	16	17	-2	-34	17	-4	-15	100	21	33	-30	8	2	11	-2	3	9	17
13 C-STW	2	14	0	3	13	-8	-4	9	-12	-2	4	-13	100	20	2	-14	-34	-17	11	23	28	13
14 C-PMT	0	10	-15	-13	12	-3	27	-16	14	-7	-3	-1	-4	100	4	1	29	24	23	13	20	19
15 C-DTC	5	-9	-2	-1	-13	11	-20	26	-6	17	4	-34	35	-14	100	8	7	7	-2	0	-7	6
16 C-HST	11	22	7	9	-1	-31	16	16	-8	32	52	-16	25	-6	18	100	-5	-7	-1	-44	-10	32
17 R-MISS	-23	-10	-64	-74	6	9	1	-24	5	-5	-5	28	-24	25	-32	-19	100	77	3	-5	-39	-5
18 C-MISS	-32	-21	-86	-93	4	5	1	-18	4	-4	-4	11	3	11	-1	-3	74	100	1	10	-34	3
19 PRESPRY	-12	1	-17	-10	-15	-3	-7	8	-4	1	4	-10	17	19	4	-1	-3	10	100	4	20	4
20 SITE05	-36	6	-36	-26	20	-2	30	-14	12	-3	9	-15	26	15	0	12	11	29	15	100	0	-16
21 SITE35	44	17	29	30	4	-3	13	1	13	8	8	-16	7	4	8	14	-24	-29	-3	0	100	0
22 SITE12	10	-1	-10	-13	4	-6	-27	26	-16	-18	-2	-21	28	-10	24	1	-1	23	21	-14	0	100

Table 26

Correlation Matrix Showing Extent and Pattern of Collinearities
Among Predictors for NARRATIVE LISTENING COMPREHENSION (TRAS - 5)
Year 1 above diagonal; Year 2 below diagonal

	1	2	3	4	5	6	7	8	.	10	11	12	13	14	15	16	17	18	19	20	21	22
1 SPNCATG	100	23	31	30	-4	16	8	-2	7	12	-2	-1	-5	4	-4	9	-20	-24	12	-38	41	4
2 PREVYR	48	100	21	12	-19	26	-5	-26	6	6	2	4	-9	16	3	-11	-21	-15	3	-26	17	5
3 PROGGRP	32	36	100	94	2	19	-11	-4	12	-7	9	2	-7	9	-1	-3	-67	-74	-4	-37	25	-14
4 PROGYX	29	35	93	100	7	13	-6	11	15	-10	4	-1	3	0	3	1	-63	-85	-5	-31	26	-10
5 R-QFL	-3	-5	0	-2	100	-9	-38	-9	34	-37	28	-10	29		-24	-28	[illegible]	-8	7	26	22	-25
6 R-DGI	-2	-4	13	7	26	100	-15	9	44	-24	-16	8	-8	24	1	45	1	-14	4	-33	-15	20
7 R-ETT	-5	-7	-11	-2	39	-16	100	-14	-34	28	-26	-2	11	4	36	8	-1	6	-6	4	-10	25
8 R-NST	4	5	-1	6	-37	-28	-45	100	32	7	-34	18	16	-8	5	55	2	-12	-6	-6	2	17
9 R-ADC	-5	-3	-5	-4	30	16	45	-49	100	-26	-14	25	37	28	-15	24	-1	-15	17	-8	19	8
10 R-SNT	7	3	-3	0	-7	-3	18	15	17	100	-13	-7	1	-2	23	2	-1	10	-1	-7	-2	2
11 R-CNT	1	6	-6	0	-8	-45	22	35	-3	32	100	-17	-19	-23	-17	-25	1	-4	2	-4	-7	-36
12 C-ADC	-2	-9	9	-2	18	27	-2	43	14	-8	-17	100	15	24	-12	6	-5	-1	-2	-6	-9	28
13 C-STW	3	-11	4	5	14	-1	-4	[illegible]	-11	-4	1	-8	100	9	5	-9	-19	-1	9	26	[illegible]	16
14 C-PNT	4	3	0	-3	16	16	24	-30	14	-13	-11	11	0	100	16	0	10	1	17	-10	17	25
15 C-DTC	5	-7	0	0	-11	13	-20	21	-6	17	3	-31	35	-12	100	7	0	-1	-5	-6	-6	18
16 C-NST	8	7	0	2	-3	-31	14	19	-8	32	50	-16	23	-13	16	100	3	-1	-2	-25	-9	18
17 R-MISS	-23	-21	-68	-75	1	-10	3	-5	4	-1	3	13	26	12	-31	-10	100	80	2	3	-33	-13
18 C-MISS	-30	-37	-86	-93	1	-9	2	-4	4	-1	2	1	1	2	-2	3	76	100	2	16	-28	-5
19 PRESFRV	-9	-9	-14	-9	-12	-2	-6	6	-2	3	3	-10	17	19	4	1	-3	9	100	1	17	3
20 SITE05	-39	-30	-36	-27	16	-12	29	-2	11	-1	11	-20	2	5	-1	14	15	29	12	100	-1	-1
21 SITE35	43	38	26	28	4	-1	13	0	13	7	6	-14	6	3	9	11	-22	-28	-2	0	100	0
22 SITE12	2	-13	-13	-14	1	-16	-25	32	-15	-15	2	-26	25	-16	22	4	2	22	18	-1	0	100

Table 27

Correlation Matrix Showing Extent and Pattern of Collinearities
Among Predictors for EXPOSITORY LISTENING COMPREHENSION (IRAS - S)
Year 1 above diagonal; Year 2 below diagonal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 SPHCATG	100	17	35	32	-2	5	17	9	8	19	-8	-7	-1	-5	-4	13	-22	-26	11	-34	48	7
2 PREVYR	43	100	-5	-9	-7	0	4	-7	7	12	-7	-11	3	-2	0	0	3	10	7	2	19	-6
3 PROGGRP	34	32	100	94	29	-4	-1	22	16	-13	-3	-12	15	-17	-6	5	-61	-77	0	-32	37	-20
4 PROGYx	30	32	93	100	32	-3	-5	31	19	-19	-3	-12	19	-21	-4	6	-68	-87	-3	-28	32	-17
5 R-OFL	-6	-8	-6	-6	100	-7	-40	-3	42	-23	-3	16	39	14	-8	-41	-23	-35	10	21	29	10
6 R-DGI	-11	-15	-6	-7	22	100	1	47	50	-21	-44	1	5	17	10	65	24	4	1	-29	-18	29
7 R-ETT	-5	3	-10	-1	39	-10	100	-44	-40	26	21	-30	-10	-18	13	3	-9	4	-3	1	-14	-1
8 R-HST	12	-1	19	21	-36	-10	-52	100	45	-23	-26	29	6	-1	-2	70	-19	-34	-7	-31	1	33
9 R-ADC	-7	-2	-6	-4	28	19	45	-56	100	-31	-35	29	41	35	-9	29	-1	-20	18	-12	20	15
10 R-SMT	8	7	1	3	-9	2	17	12	15	100	10	-19	-11	-14	15	0	4	20	-2	-4	-1	-13
11 R-CNT	6	13	2	6	-4	-43	23	32	-2	34	100	9	-17	-19	10	-27	11	3	-4	5	-9	-20
12 C-ADC	-7	-9	-4	-12	16	17	-2	-34	17	-4	-15	100	21	33	-30	8	2	11	-2	3	9	17
13 C-STW	2	-12	0	3	13	-8	-4	9	-12	-2	4	-13	100	20	2	-14	-34	-17	11	23	28	13
14 C-PMT	0	8	-15	-13	12	-3	27	-16	14	-7	-3	-1	-4	100	4	1	29	24	23	13	20	19
15 C-DTC	5	-17	-2	-1	-13	11	-20	26	-6	17	4	-34	35	-14	100	8	7	7	-2	0	-7	6
16 C-HST	11	21	7	9	-1	-31	16	16	-8	32	52	-16	25	-6	18	100	-5	-7	-1	-44	-10	32
17 R-MISS	-23	-14	-64	-74	6	9	1	-24	5	-5	-5	28	-24	25	32	-19	100	77	3	-5	-39	-5
18 C-MISS	-32	-31	-86	-93	4	5	1	-18	4	-4	-4	11	3	11	-1	-3	74	100	1	10	-34	3
19 PRESFRY	-12	-7	-17	-10	-15	-3	-7	8	-4	1	4	-10	17	19	4	-1	-3	10	100	4	20	4
20 SITE05	-36	-20	-36	-26	20	-2	30	-14	12	-3	9	-15	26	15	0	12	11	29	15	100	0	-16
21 SITE35	44	39	29	30	4	-3	13	1	13	8	8	-16	7	4	8	14	-24	-29	-3	0	100	0
22 SITE12	10	-13	-10	-13	4	-6	-27	26	-16	-18	-2	-21	28	-10	24	1	-1	23	21	-14	0	100

are correlated with entry language category at a value of roughly .45. The relation of entry language category to the various decoding tasks is lower (.35 or less), and the relation with Reading Comprehension is negligible (.25 or less).

Entry language category is positively related to the program factors in both Year 1 and 2; above-median rating in Spanish is more likely to be associated with assignment to Spanish reading instruction, with correlations of about .3. Also noticeable in both years is a pattern relating Spanish entry language category to the site contrasts. Above-median ratings in entry Spanish are more likely in the border than the nonborder sites, and in the more urban of the two nonborder sites. The import of this finding is that the teachers were apparently not totally influenced by local context, but were able to sustain a rating scale with some degree of constancy. Otherwise teachers within each locale would classify students' competence in a language relative to the prevailing level of competence within that locale, and the correlations of language rating with site contrasts would wash out.

Previous achievement has a pattern complementary to that found for English IRAS, with a few variations. Positive deviations on Spanish IRAS are positively correlated with program assignment; better performance in Spanish means that a student is more likely to be assigned to Spanish reading instruction. This relation is relatively weak in Year 1, compared with the corresponding effect in English (correlations are generally .2 or less); in Year 2, interestingly, it is again the IRAS-S measures of spoken formal language (Vocabulary Definition, and Narrative and Expository Listening Comprehension) that are most highly correlated to program assignment to Spanish reading instruction, not decoding skill in Spanish.

Previous year's achievement in Spanish is more highly related to site contrasts than was true for English. The pattern in Year 1 is such that, as one might expect, positive Spanish deviations are more likely at border than nonborder sites; remember that, for Year 1, previous year's achievement is based on a contrast of alphabet knowledge from the Stanford Foundation Skills Test. In Year 2, previous achievement (now differentiated according to IRAS components) is more closely related to the contrast between the two nonborder sites. Students from the more urban site leave first grade with higher Spanish scores in most of the IRAS components (Reading Comprehension is least affected).

Program predictors. The patterns here are identical to those observed for the English IRAS. The only noticeable change is that PROG is more closely related to MISSING, an expected finding given that Spanish instruction was rare except in those classes designated accordingly, whereas the converse was less true (instruction in English was observed on occasion in classes designated as Spanish or bilingual).

Instructional predictors. Of the 132 correlations among RAMOS and Checklist factors, 26 (20%) equalled or exceeded a criterion of .3, about the same as for the English IRAS. No consistent regularities emerge across the two years, however, with changes in magnitude and sign more typical than not. The largest correlation in the table (.65) is between NST (for RAMOS) and NST (for Checklists), both of which are indices of the number of students in an instructional group; in Year 1 there was a close correspondence between what was observed and what teachers planned.

Instruction at different sites. A quick scan of the submatrix showing the correlations between instruction and site contrasts reveals that Spanish reading instruction was much less clearly distinguished between sites than was English instruction. In Table 10, 22 of the 72 correlations in the submatrix equal or exceed .3; in Table 19, only 3 of the correlations exceed this criterion. The biggest difference between sites was the presence or absence of a systematic program of Spanish reading instruction. Differences between sites in the character of that program were negligible.

General Conclusions

The presence of substantial multicollinearity in a set of predictors can seriously jeopardize the interpretation of a regression analysis. The data structures to be analyzed, as discussed in the following section, are not free of this threat. On the other hand, the correlations presented in Tables 10 to 27 are generally modest and the more substantial values are within restricted subsets of the predictors.

Certain steps could have been taken to reduce the degree of multicollinearity. Certain factors (PROG and MISSING) could have been combined into a single indicator, reducing collinearity and increasing degrees of freedom. Nested comparisons could have led to a reduction in collinearity. The implication of this analysis is that the findings to be reported may be viewed as somewhat conservative in nature; a better controlled analysis would probably yield clearer confirmation of the relations because the presence of what amounts to random noise generally has the effect of muddying the water.

The analysis of collinearities has also pointed out some substantively important effects that merit more detailed attention. In particular, the relation between achievement and program assignment, including changes in this pattern over years, and the relations between instruction and site -- both of these relations have been reviewed in only cursory fashion in this section, but might prove quite informative.

INTEGRATIVE ANALYSES

Overview

The preceding sections have laid the groundwork for the analyses to be described below, in which the predictor factors will be employed to account for observed variability in the IRAS deviations for each of the four years of data from the various cohorts. The section begins with a presentation of zero-order and partial correlations between the predictors and the deviations. These data allow the reader to assess the relational structure at the the beginning of the regression procedure, and after the effects of the precursors (entry language category and previous year's achievement level) have been entered into the regression equation. This information is important, given the presence of multiple collinearities in the predictors.

The second part of the section presents the results of the multiple regression analyses of each of the IRAS components, for each of the four years of data and for both English and Spanish. As will be emphasized later in the section, the purpose of these analyses is descriptive rather than inferential. That is, the primary aim is not to establish statistical significance against some hypothetical population, but to determine the degree to which the various predictors in the equation contribute separably to individual variations in IRAS performance. Some investigations in recent years have led to the conclusion that school achievement is totally determined by the student's home background and other individual precursors, and that variations in school and classroom factors contribute negligibly to variations in growth. These generalizations have been undercut by more careful examinations of student performance and of classroom variables, and it is now possible to point to research showing that factors related to the conditions of instruction do "make a difference." Nonetheless, the prevailing tendency still seems to place a great deal of weight on the importance of the level of skill and knowledge with which the student enters school as a kindergartner. The regression analyses will serve to inform this question in the case of children from bilingual backgrounds.

Certain caveats warrant repetition at this point. First, the programs that we observed varied considerably in their character and extent; relatively few of the students were assigned to Spanish reading instruction in the study sites, and the relation between English and Spanish programs tended to be somewhat haphazard. Second, the decision to truncate the study at the end of second grade for the third and major cohort means that longitudinal comparisons must be made with care after the second instructional year data; the data for Years 3 and 4 are limited to a distinctive subset of the sites, those located at the border.

Finally, in this section we will reorder the IRAS components. In previous volumes and sections, the order of administration has been used throughout. This was how the data were gathered and analyzed, and it seemed most natural. Certain of the components are closer than others linguistically and in the pattern of results, and so we will now resort to the grouping mentioned earlier in passing, and repeated here for the convenience of the reader:

ORAL LANGUAGE	WORD/SENTENCE DECODING	READING COMPREHENSION
• Vocabulary Definition • Narrative Listening • Expository Listening	• Vocabulary Decoding • Letter-sound Decoding • Letter-sound Spelling • Sentence Reading	• Narrative Reading • Expository Reading

This organization has helped us understand a number of consistent trends in the data, and works quite well in most instances. The status of Vocabulary Decoding and Sentence Reading are somewhat uncertain; the former often entails both "sight word" recognition as well as decoding skills, and the latter places demands on fluency with a meaningful word string in addition to the ability to decode or recognize a word in isolation. With these cautions, we will rely on this framework for ordering the IRAS measures and for organizing the discussion in the remainder of the volume.

Correlations Between Predictor Variables and IRAS Deviates

As an introduction to the presentation of the regression analyses proper, the correlations between the IRAS deviates and each of the predictor variables are discussed next. Given the strong contributions of the variables indexing entry oral language skill and previous year's achievement (as detailed below), both the zero-order correlations and partial correlations after the effects of these two variables have been removed will be presented.

English Correlations

Tables 28 through 36 present the relevant correlations for the nine IRAS English scales. These will be discussed in the order they were entered into the regression analyses: (a) entry skills, including English oral language skill at kindergarten entry and the index of the previous year's performance (i.e., skill in the appropriate task prior to entry into the relevant instructional year), (b) nominal Spanish reading program, including the number of years enrolled in Spanish reading and the indicator of enrollment in Spanish reading in the relevant instructional year, (c) indices of instruction, including the observed instruction factors, the planned instruction factors, and the indices of missing (i.e., estimated) data for the observed and planned instruction factors, (d) attendance, and (e) site contrasts.

Table 28

Interactive Reading Assessment System - English
Correlation Summary

VOCABULARY DECODING

Variable	Year 1 r	Year 1 p/r	Year 2 r	Year 2 p/r	Year 3 r	Year 3 p/r	Year 4 r	Year 4 p/r
ENGCATG	38		40		47		42	
PREVYR	50		72		81		83	
PROGGRP	-27	-9	-31	-13	-41	-5	-36	-8
PROGYx	-33	-15	-36	-15	-36	-1	-24	-5
R-ETT	-13	-18	14	14	-6	-1	-1	-13
R-DGI	4	1	-17	-23	-23	-12	-34	-22
R-QFL	14	14	-7	-6	20	8	-16	7
R-ADC	2	1	-4	-5	-38	-15	-13	-3
R-PRD	21	25	20	-1	28	2	31	1
R-SMT	2	3	3	-7	2	10	13	15
R-NST	[illegible]	-3	-4	-1	21	37	16	28
C-ACM	-1	-1	50	25	35	-14	44	11
C-QFL	27	20	32	21	-12	13	7	2
C-STW	5	-3	1	-10	33	6	21	20
C-PMT	41	35	3	-15	13	15	-46	-13
C-GRV	-6	-4	-14	-5	-2	28	-16	-9
R-MISS	-12	-6	-28	-14	-4	1	10	22
C-MISS	-25	-13	-30	-15	-8	10	2	3
PRESPRY	24	24	12	9	16	3	13	10
SITE05	-14	-10	0	15	-	-	-	-
SITE35	17	23	7	-8	-	-	-	-
SITE12	-8	-13	-14	-14	-11	7	-11	-1

921

Table 29

Interactive Reading Assessment System - English
Correlation Summary

VOCABULARY DEFINITION

Variable	Year 1 r	Year 1 p/r	Year 2 r	Year 2 p/r	Year 3 r	Year 3 p/r	Year 4 r	Year 4 p/r
ENGCATG	46		37		35		44	
PREVYR	36		52		69		49	
PROGGRP	-45	-31	-31	-6	-30	-9	-20	22
PROGYx	-50	-34	-31	-7	-25	-6	0	10
R-ETT	8	5	16	15	-7	-8	-3	-7
R-DGI	-2	-5	-14	-7	-8	14	-26	-17
R-QFL	15	12	4	8	27	13	-11	4
R-ADC	-11	-14	6	30	-24	2	-16	-16
R-PRD	2	3	18	14	30	23	19	-2
R-SMT	-20	-20	-15	-16	10	11	9	18
R-NST	-1	-4	-5	-12	11	22	17	33
C-ACM	-20	-19	46	36	30	2	52	37
C-QFL	31	23	33	24	-21	-6	10	9
C-STW	-3	-8	-1	-9	42	21	1	-16
C-PMT	32	21	8	11	22	21	-33	-14
C-GRV	8	10	-15	-13	-7	16	-15	-4
R-MISS	-20	-15	-17	-1	-8	5	17	23
C-MISS	-33	-21	-20	-5	-8	5	10	8
PRESPRY	3	-1	8	11	21	5	15	13
SITE05	3	5	13	12	-	-	-	-
SITE35	-10	-8	-5	-1	-	-	-	-
SITE12	-8	-11	-25	-25	-28	-14	-12	3

922

Table 30

Interactive Reading Assessment System - English
Correlation Summary

LETTER-SOUND DECODING

Variable	Year 1 r	Year 1 p/r	Year 2 r	Year 2 p/r	Year 3 r	Year 3 p/r	Year 4 r	Year 4 p/r
ENGCATG	31		28		44		44	
PREVYR	36		57		68		87	
PROGGRP	-18	-19	-29	-21	-39	-7	-28	15
PROGYx	-21	-4	-29	-19	-36	-7	-25	3
R-ETT	13	-18	1[illegible]	16	-3	6	16	8
R-DGI	-8	-11	-2[illegible]	-23	-28	-5	-38	-22
R-QFL	16	16	-8	-2	11	1	4	20
R-ADC	-10	-11	1	5	-38	-17	-6	-22
R-PRD	10	10	3	-8	25	22	20	-13
R-SMT	6	7	2	-1	2	7	13	18
R-NST	-7	-10	8	7	6	-8	9	-8
C-ACM	-6	-5	43	31	25	7	32	8
C-QFL	21	15	27	14	-3	18	-11	-3[illegible]
C-STW	4	-1	6	6	27	11	27	9
C-PMT	30	23	-7	-14	14	20	-37	14
C-GRV	6	8	-7	-10	-4	-2	-18	-23
R-MISS	0	5	-16	-6	4	18	7	1
C-MISS	-14	-2	-19	-8	-2	11	2	[illegible]
PRESPRY	19	18	9	10	20	8	4	-20
SITE05	6	11	12	10	-	-	-	-
SITE35	30	35	14	-1	-	-	-	-
SITE12	-21	-25	-12	1	-7	7	-11	-20

323

Table 31

Interactive Reading Assessment System - English
Correlation Summary

LETTER-SOUND SPELLING

Variable	Year 1 r	Year 1 p/r	Year 2 r	Year 2 p/r	Year 3 r	Year 3 p/r	Year 4 r	Year 4 p/r
ENGCATG	29		37		44		32	
PREVYR	35		70		78		71	
PROGGRP	-21	-7	-30	-15	-39	-12	-38	-17
PROGYx	-27	-13	-33	-15	-32	-4	-17	-11
R-ETT	-14	-17	11	2	-7	-8	3	2
R-DGI	-1	-4	-17	-16	-27	-16	-41	-14
R-QFL	18	17	-4	6	14	3	-17	2
R-ADC	2	1	-4	1	-32	-12	6	9
R-PRD	13	15	13	-3	15	1	26	17
R-SMT	6	7	4	-4	13	21	6	-1
R-NST	1	-2	7	11	-1	-8	11	23
C-ACM	-2	-1	38	22	34	4	30	-9
C-QFL	23	17	28	16	-14	1	6	11
C-STW	5	-1	1	-3	41	29	17	-3
C-PMT	33	27	4	-8	15	15	-42	-16
C-GRV	4	7	-5	-7	-16	-10	-18	-7
R-MISS	-2	3	-28	-14	-23	-8	-9	-30
C-MISS	-16	-6	-28	-15	-17	-2	-12	-18
PRESPRY	17	16	12	10	29	23	14	6
SITE05	-	4	8	9	-	-	-	-
SITE35	30	34	8	-15	-	-	-	-
SITE12	-16	-20	-16	-7	-26	-15	-16	-1

Table 32

Interactive Reading Assessment System - English
Correlation Summary

SENTENCE READING

Variable	Year 1 r	Year 1 p/r	Year 2 r	Year 2 p/r	Year 3 r	Year 3 p/r	Year 4 r	Year 4 p/r
ENGCATG	35		39		53		40	
PREVYR	53		60		76		87	
PROGGRP	-23	-2	-47	-41	-45	-1	-34	-13
PROGYx	-29	-10	-49	-38	-37	1	-30	-25
R-ETT	-6	-9	27	32	1	9	-12	-10
R-DGI	-3	-2	-20	-25	-25	13	-17	-2
R-QFL	13	8	7	12	9	-5	-24	-2
R-ADC	-10	-4	-2	9	-40	-6	-25	-9
R-PRD	10	14	4	-13	37	16	36	15
R-SMT	-7	-6	-9	-22	15	29	-6	-1
R-NST	10	12	17	11	3	2	22	16
C-ACM	27	26	47	35	40	10	44	-4
C-QFL	22	21	28	21	0	34	-2	-11
C-STW	31	27	22	20	15	-1	16	4
C-PMT	38	29	-7	-17	0	21	-43	-18
C-GRV	-20	-19	-11	-13	-3	18	15	17
R-MISS	-6	-1	-32	-21	-3	10	-9	-12
C-MISS	-23	-10	-37	-26	-12	5	-7	-13
PRESPRY	20	18	17	18	20	12	22	7
SITE05	-15	-15	10	18	-	-	-	-
SITE35	7	12	-12	-18	-	-	-	-
SITE12	8	10	9	8	21	20	21	7

925

Table 33

Interactive Reading Assessment System - English
Correlation Summary

NARRATIVE READING COMPREHENSION

Variable	Year 1 r	Year 1 p/r	Year 2 r	Year 2 p/r	Year 3 r	Year 3 p/r	Year 4 r	Year 4 p/r
ENGCATG	37		39		50		36	
PREVYR	49		60		76		79	
PROGGRP	-25	-7	-42	-30	-50	-19	-28	13
PROGYR	-31	-13	-45	-31	-42	-15	-26	-13
R-ETT	-4	-7	22	25	2	-1	12	13
R-DGI	0	-4	-21	-28	-18	10	-29	-15
R-QFL	10	9	4	3	12	-3	-13	9
R-ADC	-4	-6	-3	-1	-38	-6	-9	12
R-PRD	14	17	9	-10	40	15	31	6
R-SMT	-3	-4	-5	-11	2	3	4	1
R-NST	7	4	10	10	9	16	16	30
C-ACM	6	8	48	31	35	-5	42	7
C-QFL	25	18	28	17	-8	13	5	-8
C-STW	10	3	21	21	19	2	18	1
C-PMT	30	21	-5	-16	-2	11	-53	-38
C-GRV	-9	-7	-10	-5	0	21	4	9
R-MISS	-12	-6	-34	-26	-1	12	9	12
C-MISS	-23	-10	-35	-27	-12	7	11	19
PRESPRY	17	16	19	17	21	11	9	-19
SITE05	-10	-5	11	18	-	-	-	-
SITE35	4	7	-6	-8	-	-	-	-
SITE12	1	-2	7	10	13	7	8	-2

Table 34

Interactive Reading Assessment System - English
Correlation Summary

EXPOSITORY READING COMPREHENSION

Variable	Year 1 r	Year 1 p/r	Year 2 r	Year 2 p/r	Year 3 r	Year 3 p/r	Year 4 r	Year 4 p/r
ENGCATG	29		37		53		37	
PREVYR	48		47		66		76	
PROGGRP	-20	-3	-45	-37	-50	-21	-31	17
PROGYx	-23	-7	-46	-36	-45	-21	-30	-27
R-ETT	1	-2	25	26	4	5	10	15
R-DGI	-9	-9	-27	-33	-20	6	-28	-14
R-QFL	16	12	5	1	14	-1	-16	1
R-ADC	-10	-6	-1	7	-39	-12	-15	-1
R-PRD	2	4	3	-5	41	19	34	15
R-SMT	-8	-7	-4	-6	0	7	9	10
R-NST	3	2	15	10	1	-1	10	13
C-ACM	26	26	47	39	43	19	40	1
C-QFL	20	18	27	21	-11	12	5	-5
C-ST[illegible]	22	16	29	29	21	13	23	10
C-[illegible]MT	24	15	-9	-13	-6	8	-49	-32
C-GRV	-7	-5	-8	-8	-10	8	1	7
R-MISS	2	3	-31	-23	-3	9	14	19
C-MISS	-17	-5	-32	-24	-14	1	10	17
PRESPRY	11	7	19	20	17	7	8	-17
SITE05	3	5	14	13	-	-	-	-
SITE35	4	7	-2	-1	-	-	-	-
SITE12	-3	4	10	14	11	3	10	8

927

Table 35

Interactive Reading Assessment System - English
Correlation Summary

NARRATIVE LISTENING COMPREHENSION

Variable	Year 1 r	Year 1 p/r	Year 2 r	Year 2 p/r	Year 3 r	Year 3 p/r	Year 4 r	Year 4 p/r
ENGCATG	47		47		47		36	
PREVYR	21		63		79		74	
PROGGRP	-39	-25	-46	-26	-35	5	-11	16
PROGYx	-39	-22	-44	-26	-34	11	-10	-8
R-ETT	14	11	21	19	3	10	9	8
R-DGI	-18	-22	-24	-15	-23	9	-25	4
R-QFL	18	14	2	11	20	4	-10	8
R-ADC	-17	-20	1	5	-29	2	-13	-1
R-PRD	-13	-13	6	4	29	6	23	7
R-SMT	-20	-18	-19	-17	5	2	8	3
R-NST	1	-1	7	-4	-10	-4	24	43
C-ACM	-21	-17	41	34	30	18	50	23
C-QFL	24	16	32	20	-16	8	-12	-21
C-STW	6	6	11	10	28	7	13	12
C-FMT	20	7	-7	-6	7	13	-39	-32
C-GRV	10	9	-10	-8	-10	16	-6	7
R-MISS	-6	1	-24	-13	-6	8	3	7
C-MISS	-19	-8	-32	-23	-14	4	9	12
PRESPRY	3	1	7	11	14	4	0	-12
SITE05	25	27	23	13	-	-	-	-
SITE35	-10	-6	-8	-3	-	-	-	-
SITE12	-10	-10	-10	-4	-4	7	6	17

Table 36

Interactive Reading Assessment System - English
Correlation Summary

EXPOSITORY LISTENING COMPREHENSION

Variable	Year 1 r	Year 1 p/r	Year 2 r	Year 2 p/r	Year 3 r	Year 3 p/r	Year 4 r	Year 4 p/r
ENGCATG	49		41		51		34	
PREVYR	31		48		67		74	
PROGGRP	-39	-21	49	-37	-38	1	-14	11
PROGYx	-41	-22	-48	-36	-37	-1	-11	-11
R-ETT	7	3	25	25	1	7	14	21
R-DGI	-8	-6	-26	-28	-27	-4	-31	-8
R-QFL	24	18	-0	-2	15	-2	-7	8
R-ADC	-14	-11	6	13	-30	-1	-20	-23
R-PRD	2	7	6	-1	31	5	20	4
R-SMT	-14	-9	-16	-19	3	8	22	28
R-NST	-1	-1	14	7	-5	-1	12	15
C-ACM	3	1	41	38	31	7	47	22
C-QFL	20	16	34	28	-14	13	-16	-31
C-STW	20	13	16	15	28	18	19	19
C-PMT	21	10	-10	-16	8	23	-36	-22
C-GRV	-4	-1	-7	-6	-9	19	-10	3
R-MISS	1	4	-31	-19	-3	12	10	16
C-MISS	-30	-18	-37	-24	-11	8	8	13
PRESPRY	-1	-3	5	9	16	2	-	-17
SITE05	0	-2	25	27	-	-	-	-
SITE35	-5	-1	-6	-2	-	-	-	-
SITE12	-7	-8	-7	-6	-3	2	3	14

923

Entry Skills

For English entry level, the first variable entered in the regression analyses, the correlations are generally uniform across scales and stable across instructional years within scales. Recall that the kindergarten teacher rated the child in both English and Spanish on entry to school, and these ratings were then used to categorize the student as relatively high or low. Correlations of this measure with IRAS-E deviations did differ somewhat for different IRAS components, ranging from .35 to .50 for the oral language components, with a mean of .43, and from .30 to .50 for the decoding and reading comprehension components, with a mean of .39 over all analyses; this stands in contrast with the pattern for previous achievement.

These relationships suggest that relatively high English oral language skill at kindergarten entry is positively associated with above-average performance in each of the IRAS-E component skills assessed throughout the early grades. Thus, for the bilingual sample, students who begin school with relatively greater English oral skills are better able to profit from the English reading instruction received, and further, the advantages in acquired skills continue to be evidenced through fourth grade exit. Note that these students are also more likely to receive different instructional programs than those whose entry level in English is judged to be relatively low, as will be seen below and as noted in the previous section.

Knowledge of the English alphabet was employed as an estimate of first-grade entry literacy skill. It is consistently correlated with each of the first-year deviates. The correlation coefficients average about .4, with higher coefficients (about .45) for decoding based components (Vocabulary Decoding, Sentence Reading, and Reading Comprehension) and lower values (about .3) for oral language based components (Vocabulary Definitions and Listening Comprehension). Thus, knowledge of the alphabet at kindergarten entry is associated with literacy skill at the end of first grade, more strongly for reading (i.e., print) components than for oral language components, a pattern opposite to that for the entry language category measure discussed above.

For Years 2 through 4, the corresponding IRAS index from the assessment during the preceding Spring was entered as the measure of previous year's achievement. The correlation pattern for these indices is similar across IRAS scales, generally showing increases in magnitude from the second instructional year (about .6) to the fourth instructional year (about .8). The trend is less strong over years for the oral language components (from .55 in Year 2 to .65 in Year 4) due to ceiling effects (recall that for students in the aggregate, both Vaocabulary Definition and Listening Comprehension scores were remarkably high at the beginning of first grade). For the Decoding and Reading Comprehension measures, the increase from Year 2 to Year 4 in the correlations was much more striking (.60 to .80).

This pattern means that the relative standing of individual students becomes more fixed in place with increased schooling. That is, in the early grades, the degree of growth is not as dependent upon entry skill as in the latter grades. This pattern was remarked upon in the previous section, and will appear in subsequent analyses, along with a discussion of possible interpretations for the finding.

Nominal Reading Program Indicators

In the third regression step, the two general Spanish reading program indicators were entered: PROGGRP, the t[illegible] l number of years of enrollment in Spanish reading (ranging from 0 [illegible] 5), and PROGYx, the dichotomous indicator of enrollment in Spanish reading in the instructional year x currently under analysis (0 for no current enrollment and +1 for enrollment). Recall that these two factors are highly correlated, especially in Years 1 and 2.

The zero-order correlations of the program variables with the IRAS scale deviates are negative, generally ranging from -.3 to -.5 over scales and years. Because of variations in cohorts that are confounded with variations in sites, changes in these variables over years need to be interpreted with care, especially those for Years 3 and 4.

A trend that appears in Years 1 and 2 is for the program variables to have the highest zero-order correlations with the oral language IRAS scales (-.40 to -.50), and the lowest correlations with the reading comprehension scales (smaller in absolute value than -.30). At the zero-order level, then, assignment to Spanish reading instruction (from both the longitudinal and current year points of view) is negatively related to relative growth in all IRAS-E components, a relation accounting for 10 to 25 percent of the variance in the IRAS deviations.

To be sure, assignment to Spanish reading instruction reflects the student's precursor standing, and so we would expect a change in these patterns when precursors are taken into account. The partial correlations of the program factors with IRAS deviations are indeed reduced compared with the the zero-order correlations, especially for Years 3 and 4. The partial correlations remain consistently negative and rather robust in Years 1 and 2, however, especially for the oral language components of IRAS, where they range from -.20 to almost -.40. (The small partials for Vocabulary Definition in Year 2 are a somewhat surprising departure from this generalization.) Partial correlations with decoding scales are generally small, but are modest for Sentence Reading and Reading Comprehension, especially in Year 2.

In summary, enrollment in Spanish reading programs is negatively related to the set of English reading deviates, though the strength of the relationship is reduced once English entry skill and previous year's performance are considered. The independent effect of assignment to t[illegible]e programs is greatest in Year 2, especially for those IRAS scales designed to assess the development of skills in handling spoken English in formal settings. The pattern of partial correlations in Years 3 and 4, with a restricted sample of students, is less consistent. One faint but intriguing trend in the data for Year 4 does merit comme . The partial correlations of PROGGRP are positive for six of the nine IRAS scales, suggesting that, once the precursor effects are taken into account, students who have been assigned to more years of Spanish reading instruction may begin to benefit somewhat during the fourth year of instruction. To repeat, this trend is very slight, and is restricted to the data from the border sites.

Instruction

Two sets of instructional indices are entered on the next step of the regression, first those obtained from the observation data (RAMOS factor scores), then those derived from the teacher interviews concerning planned instruction (Checklist factor scores). The discussion of the zero-order and partial correlations of these predictors will focus on consistent patterns, especially those found in the data for the first two instructional years.

While the partial correlations showed substantial differences from the zero-order correlations for the two program variables discussed above, such differences were in general much smaller for the correlations concerning the instructional dimensions. Such might be expected given that program assignments partly depend upon oral language entry skill, while the structure of instruction within distinct segregated programs likely depends more upon other non-linguistic factors. Some changes are observed, and it should be reiterated that the precursors are correlated with program assignments. Each of the instructional dimensions is discussed separately below, observed instruction first, then planned instruction.

Observed instruction. The first RAMOS factor, engaged text time (ETT), is an index of the emphasis placed on engaged reading with text. In the correlations displayed in Tables 28 through 36, this factor shows small negative relationships in the first instructional year for the set of decoding assessments (Vocabulary Decoding, Letter-sound Decoding, and Letter-sound Spelling), with the partial correlations averaging about -.2. Conversely, in the second instructional year, positive relationships are found, small for decoding (about .15), and larger (at about .25) for those assessments concerned with connected text (Sentence Reading, Narrative and Expository Reading and Listening Comprehension). For the latter indices, positive correlations are also found in Years 3 and 4, but they are of lesser magnitude. Overall, increased reliance on text materials seems to delay the acquisition of decoding skills in first grade where such skill is initially developed. However, once some initial success in decoding is found, subsequent involvement of students in text material seems essential for the advancement of reading skill, more so for comprehension than for decoding.

The second RAMOS factor is an index of direct instruction delivered by the teacher to groups of students rather than individuals (direct group instruction, or DGI). For the first instructional year, no systematic trends appear, but for the second instructional year, the coefficients are negative, averaging about -.2 over the nine IRAS scales, with only modest changes between the zero-order and partial coefficients. (Neither ETT nor DGI are apparently related to program assignment in Years 1 or 2.) For the third and fourth instructional years, the zero-order coefficients are generally smaller than Year 2, and are often reduced to values near zero once the contributions of entry skill levels are removed. Where there is an effect, it is generally negative and generally associated with decoding scales.

Direct group instruction does not show systematic effects in the initial year of instruction, but has a negative influence across the set of reading skills assessed in each of the subsequent years, most noticeably in the second instructional year. One interpretation of this finding is that even though such instruction represents increased time spent with an instructor providing increased amounts of direct instruction, the accompanying practice of providing such to large groups of students counters the expected positive result.

The third factor, quality of formal language (QFL), is a measure of the formal language demands made upon the students. This factor consistently shows a small positive relationship to each of the IRAS deviates in the first instructional year (averaging about .15), with partial correlations only slightly reduced from the zero-order values. There is no consistent pattern relating this index with the deviations for Year 2 and following years.

Thus, as formal language demands increase, there is a slight tendency toward increased skill during Year 1 in each of the reading components assessed. The relationship most likely reflects the contribution of explicit instruction on letter-sound correspondences during this year. The latter is one of the few curriculum areas that appears with any discernible frequency as a "formal language demand."

The amount of decoding (ADC), the next RAMOS factor, is an index of the relative amount of time devoted to instruction in decoding. The only noticeable effects appear in Year 1, where the zero-order and partial correlations are slightly negative with the oral language scales in IRAS; an emphasis on decoding is associated with less growth in spoken English skills. A secondary pattern appears in Years 3 and 4, the former more clearly than the latter, in which ADC is negatively correlated with IRAS decoding and reading comprehension scales. It appears that relatively greater emphasis in these grades on decoding, presumably for students who are having problems in reading English, is counterproductive. You may recall from the descriptive account in Volume 6 that such decoding instruction was generally non-explicit.

The fifth RAMOS factor provides a measure of the conditions promoting high individual student productivity (PRD). In the first instructional year, this factor shows small positive relationships (partial correlations averaging about .15) with those IRAS scales most directly related to decoding (Vocabulary Decoding, Letter-sound Decoding, Letter-sound Spelling, Sentence Reading, and Narrative Reading Comprehension), but not with the assessments most directly tapping oral language skill (Vocabulary Definitions and Listening Comprehension). The relationships within the second instructional year are negligible. In the third and fourth years, the zero-order coefficients tend to be modestly positive (about .28 on the average), but are sharply reduced when the precursors (and hence the program factors) are partialed out. The pattern in the last two years is consistent with a situation in which the more able students become noticeably more productive in academic tasks than those rated less able and performing more poorly -- and the relatively lower productivity is associated through this mediator with poorer year-end performance.

The next factor provides a measure of the relative usage of secondary materials (SMT), though this interpretation is somewhat oversimplified. The most prominent finding is that the usage of secondary materials is negatively related to oral language deviates in the early grades even after oral language entry has been taken into account. This suggests that when such supplementary materials are employed, they tend to be associated with a lowered emphasis on formal language, an interpretation supported by the negative correlation found between this factor and that assessing the quality of formal language (QFL).

The final factor reflects the number of students constituting an instructional group (NST). The correlations are in general negligible except in the fourth instructional year where small positive partials are found, averaging about .25, for each of the IRAS scales except Letter-sound Decoding. Instructional group size is not related to literacy skill acquisition in these data.

In summary, the correlation pattern for the seven instructional factors based on classroom observations, reveals the following. First, relatively superior literacy skills in general were associated with increased engaged text time (especially for comprehension skills) and increased formal language demands in the instruction provided. The amount of direct group instruction, while allowing for increased amounts of direct instruction, was negatively related to the set of literacy skills assessed, perhaps due to the practice of providing such instruction to larger groups of students. Second, decoding skills were positively related to productivity, but negatively related to the amount of time devoted to decoding instruction, though such decoding instruction was found largely to be of low quality. Finally, oral language skills are negatively related to the use of secondary materials, but such usage was associated with a lowered emphasis on formal language.

Planned instruction. Turning to those dimensions of instruction based on teacher plans, the first Checklist factor provides a measure of the relative amount of time planned for instruction in comprehension (ACM). In the first instructional year, the relations are small and inconsistent. In Year 2, the coefficients are uniformly positive across the nine IRAS scales, relatively large, and unaffected by the precursors (the partials average about .3). In Years 3 and 4, the zero-order correlations are again positive and of modest extent, but the relation is mediated by the precursor variables, so that the partial correlation coefficients are, with a few exceptions, generally near zero.

Thus, the amount of time in Year 1 planned for comprehension instruction is not consistently related to any of the IRAS scales (other research suggests that comprehension instruction is negligible in most primary classrooms). In Year 2 and subsequent years, the factor shows positive effects across the set of IRAS assessments; these are generally reduced in the latter years when entry skills are taken into account, perhaps because of the limited nature of the earliest cohorts.

The second factor, quality of formal language (QFL), constitutes an index of the formal language demands planned by the teacher. For the first two instructional years, the zero-order correlations are positive across the set of IRAS scales, with the partial correlations slightly reduced to an average of about .2. The patterns in Years 3 and 4 are weak and inconsistent.

The third Checklist factor, seatwork (STW), is an index of the relative amount of time to be devoted to independent seatwork as opposed to group work. For the set of comprehension scales, small positive coefficients across years are found with the largest values appearing in Year 2 (partial correlations averaging about .2). The only other discernible pattern for this index is found in Years 3 and 4, where there is a positive relation to IRAS scales of single-word decoding; the zero-order correlations generally exceed .30, but are often negligible when the precursors are taken into account.

Recall from the discussion of the factor inter-correlations in Volume 6 that the amount of comprehension (ACM) showed a moderate positive relationship to the amount of seatwork (STW), suggesting that for students given relatively larger amounts of seatwork, there was an associated increase in the amount of comprehension instruction. This relationship most likely contributes to part of the associations found here between seatwork and comprehension scale deviates.

The next Checklist factor, primary materials (PMT), assesses the planned usage of primary materials (although the factor was more complex than the label indicates). In Year 1 the zero-order correlations average about .30, dropping to .20 when the precursors are taken into account. The relation is strongest for the decoding scales of IRAS. The patterns in Years 2 and 3 are weak and inconsistent. In the fourth instructional year negative values are found (partial correlation averages of -.2), with the strongest relationships in the four comprehension scales. Recall that the quality of instruction devoted to sentence/text meaning (IFSTMN) negatively loaded on this factor, and thus while one may expect primary material usage to have a positive influence on the acquisition of reading, the accompanying reduction in the quality of instruction found in this data set may have led to a weakening influence. Again, recall that the sample for Year 4 is limited.

The final factor is a measure of the relative amount of time planned for group instruction in word meaning, labelled GRV for group vocabulary instruction. This factor had no consistent effects on IRAS deviations.

In summary, the correlation pattern for the five instructional factors based on teacher instructional plans, suggests the following. First, relatively higher literacy skills were associated with increased amounts of (planned) instruction entailing greater amounts of formal language demand, and with increased planned usage of primary materials. Second, relatively superior comprehension skills were associated with increased amounts of time on comprehension instruction and with increased amounts of seatwork, two factors that were themselves correlated in this data set.

Missing instruction indicators. Following the entry of the above dimensions of instruction, observed and planned, the missing data indicators for these two sets of instructional data were entered to assess whether estimated data were associated with deviations. The missing indicators were often negatively correlated with one or more of the IRAS scales during a given year. The pattern is quite uneven in Year 1. In Year 2 the zero-order correlations range between -.20 and -.35 for most scales, but are generally no more than -.20 when precursors are partialled out. The relations in Years 3 and 4 are again uneven.

A missing (and hence estimated) RAMOS or Checklist variable might arise for several reasons. First, as mentioned earlier, missing data appear in the instructional indices for a given language of instruction when a student is exclusively involved in a reading program taught in the other language. Second, some students were involved in special pullout classes (e.g., English-as-a-Second-Language) and thus left the classroom during the observed period. Finally, some migrant students were represented in the sample and were not enrolled in school at all during certain (sometimes extended) periods, such absenteeism being generally restricted to early Fall and late Spring.

Attendance

The average percentage of days in school attendance during a given instructional year x, PRESPRYx, was entered in the next regression step. The effects of this variable generally appear in the first three instructional years, and most prominently for those skills largely based in decoding (average partial correlations of .15 across the three instructional years and the six relevant scales). Simply put, literacy skills tend to show greater improvement with increased exposure to instruction. Conversely, failure to attend school means that the student does more poorly on those components of reading that are most closely associated with schooling, but there is less effect, at least in the short run, on the development of spoken English among the students in this bilingual sample.

Site Contrasts

In the final regression step, the three site contrast variables were entered: SITE05, SITE35, and SITE12. Given the cohort structure of the study, the first two site contrasts are applicable only during the first two instructional years. For the contrast between border and nonborder sites (SITE05), small positive partial correlations averaging about .15 are found in the second instructional year for the four comprehension scales (Narrative and Expository Reading and Listening Comprehension). Given the coding of this variable, comprehension skills evidenced in the second instructional year tend to be more highly developed in the two nonborder sites than in the three border sites, once differences in entry skills have been taken into account.

For the contrast between the nonborder sites (SITE35), positive partial correlations are found for the decoding scale deviates within the first instructional year, averaging about .3. This suggests that decoding skills do not differ between these two sites at the end of first grade (once precursors have been considered), but in the subsequent year, students from the more urban site possess relatively superior decoding skills to those target students from the less urban site.

For the contrast between the border sites (SITE12), no consistent effects of any magnitude were observed between the site contrast and patterns of deviations.

Summary of the English Correlations

In summary, the correlations between the regression predictor variables and English IRAS deviates reveal the following:

English kindergarten entry skill is associated with above average performance in each of the IRAS component literacy skills assessed throughout the early grades.

Entry literacy skill is similarly related to exit literacy skill, and increasingly so over grade levels, thus suggesting that students tend to become academically "locked in place" with increased schooling.

Enrollment in Spanish reading programs is generally negatively associated with acquired English literacy skill. However, there is some indication of relatively superior English literacy skills at fourth grade exit for those students with longer (longitudinal) enrollments in h Spanish reading programs (though the sample is limited in is instructional year).

Relationships for both observed and planned instructional dimensions suggest that (a) literacy skills are advanced by instruction that makes strong formal language demands on students, by instruction that employs primary materials, and by instruction that engages students in work with text materials, and (b) comprehension skills and vocabulary skills are advanced by increased amounts of instructional time devoted to such skill development; decoding skills show the opposite relationship, perhaps because of the relatively low quality of such instruction found in this data set.

Literacy skills tend to show greater improvement with increased exposure to instruction -- the more opportunity for learning, the greater the skill acquired. This relationship was in many instances not generic in the present data set. Instead, more time on a particular component was correlated with growth in that component.

Finally, some site contrasts are evident even after site differences due to entry skill have been removed, but these are relatively isolated.

Spanish Correlations

Tables 37 through 45 present the relevant correlations for the nine IRAS Spanish scales. In parallel with the previous section, these will also be discussed in the order they were entered into the regression analyses.

Entry Skills

For oral Spanish entry skill, the first variable entered in the regression analyses, the correlations, as in the English data, tend to be relatively uniform across scales and across instructional years. The relation of Spanish entry category to the oral scales of IRAS tend to be higher than for the scales that require decoding; the former average about .40 and the latter average .32, with no trend over years. These data show that relatively high Spanish oral skills at entry to kindergarten continue to be positively associated with above-average performance in Spanish literacy skill throughout the early grades, especially with skills in formal use of the spoken language.

Knowledge of the Spanish alphabet was employed as an estimate of general literacy on entry to first grade. Kindergartner's knowledge of the Spanish alphabet was typically minimal, and so this index was only weakly (though positively) related to the IRAS scales. For subsequent years, the corresponding IRAS scale deviation from the previous year served as the index of previous year's performance. The correlation pattern for these indices is similar to that for the English data: coefficients increasing from .60 in Year 2 to .70 in Year 4 for oral language scales, and from .70 in Year 2 to more than .80 in Year 4 for the single-word decoding skills. Correlation of

Table 37

Interactive Reading Assessment System - Spanish
Correlation Summary

VOCABULARY DECODING

Variable	Year 1 r	Year 1 p/r	Year 2 r	Year 2 p/r	Year 3 r	Year 3 p/r	Year 4 r	Year 4 p/r
SPNCATG	32		36		32		31	
PREVYR	8		74		83		83	
PROGGRP	14	4	24	15	-4	-1	-4	13
PROGYx	17	9	25	16	-9	-2	0	26
R-QFL	11	13	3	9	4	-4	-12	-10
R-DGI	-1	-7	-13	4	-11	-17	32	31
R-ETT	4	2	17	16	-5	3	-15	-13
R-NST	-6	-5	-14	-12	-3	5	-7	-10
R-ADC	9	7	7	15	-7	8	-4	-1
R-SMT	4	-1	9	6	4	12	-12	-8
R-CNT	8	9	-3	-11	-30	-33	-32	-31
C-ADC	-29	-30	-15	-14	-10	-4	-5	-8
C-STW	10	12	6	2	13	11	-14	1
C-PMT	-17	-20	20	8	-18	-8	4	6
C-DTC	1	2	17	12	-12	-2	1	4
C-NST	9	6	4	-20	-15	-17	2	-13
R-MISS	-16	-10	-28	-21	-11	-12	5	-31
C-MISS	-17	-10	-35	-26	4	-2	8	-24
PRESPRY	21	18	5	4	10	-1	3	-10
SITE05	-8	5	-11	-1	-	-	-	-
SITE35	27	16	38	21	-	-	-	-
SITE12	-22	-25	-22	-12	-30	-7	-35	-23

Table 38

Interactive Reading Assessment System - Spanish
Correlation Summary

VOCABULARY DEFINITION

Variable	Year 1 r	Year 1 p/r	Year 2 r	Year 2 p/r	Year 3 r	Year 3 p/r	Year 4 r	Year 4 p/r
SPNCATG	45		43		24		36	
PREVYR	18		61		61		72	
PROGGRP	38	27	36	16	13	8	17	23
PROGYx	33	24	35	18	8	5	18	23
R-QFL	3	2	2	7	11	-4	-12	-3
R-DGI	-3	-11	5	12	5	7	22	17
R-ETT	-1	1	8	10	-9	-10	-14	-4
R-NST	-20	-19	-9	-10	-11	2	-1	-9
R-ADC	-3	-8	5	13	-5	4	-7	3
R-SMT	10	6	10	10	9	14	-12	-2
R-CNT	20	18	-3	-8	-22	-10	-22	-17
C-ADC	-24	-21	-2	1	0	9	-3	4
C-STW	-9	-6	-6	-6	19	20	-13	8
C-PMT	-9	-9	14	1	-19	-10	3	17
C-DTC	0	6	8	14	-10	-1	7	10
C-NST	-6	-9	1	-6	-16	-12	8	-7
R-MISS	-24	-17	-31	-20	-26	-28	-3	-22
C-MISS	-27	-20	-45	-26	-11	-6	-11	-21
PRESPRY	13	8	-7	-2	5	-2	-13	-25
SITE05	-26	-11	-26	-8	-	-	-	-
SITE35	31	14	42	25	-	-	-	-
SITE12	-39	-38	-34	-25	-38	-13	-46	-30

Table 39

Interactive Reading Assessment System -- Spanish
Correlation Summary

LETTER-SOUND DECODING

Variable	Year 1 r	Year 1 p/r	Year 2 r	Year 2 p/r	Year 3 r	Year 3 p/r	Year 4 r	Year 4 p/r
SPNCATG	36		37		27		18	
PREVYR	8		68		71		81	
PROGGRP	14	3	23	14	-12	-7	-19	-5
PROGYx	19	9	26	17	-13	-5	-11	3
R QFL	17	20	-1	4	7	6	-8	4
R-DGI	-7	-13	-13	4	-9	-8	40	22
R-ETT	-1	-4	18	13	0	8	-13	2
R-NST	-6	-6	-8	-7	-16	-23	-16	-18
R-ADC	8	5	10	22	-1	10	2	10
R-SMT	4	-1	11	9	5	0	-9	5
R-CNT	12	13	-1	-9	-25	-14	-40	-21
C-ADC	-24	-25	-18	-9	-1	11	-8	-10
C-STW	12	14	12	9	31	36	-13	-5
C-PMT	-20	-23	15	-1	-17	-10	2	-1
C-DTC	-3	-2	18	18	-12	-10	-8	5
C-NST	0	-3	7	-13	-33	-41	-7	-12
R-MISS	-24	-18	-31	-22	-2	-1	9	-7
C-MISS	-24	-16	-34	-23	13	9	18	-6
PRESPRY	15	12	7	10	7	-9	15	6
SITE05	-5	10	-9	1	-	-	-	-
SITE35	30	18	51	40	-	-	-	-
SITE12	-23	-27	-20	-7	-32	-16	-18	5

941

Table 40

Interactive Reading Assessment System - Spanish
Correlation Summary

LETTER-SOUND SPELLING

Variable	Year 1 r	Year 1 p/r	Year 2 r	Year 2 p/r	Year 3 r	Year 3 p/r	Year 4 r	Year 4 p/r
SPNCATG	33		41		28		35	
PREVYR	18		73		80		81	
PROGGRP	17	5	24	11	5	17	0	-6
PROGYx	16	6	28	19	2	13	3	1
R-QFL	14	18	3	[illegible]	5	-4	-10	2
R-DGI	2	-6	-6	5	-9	-20	26	4
R-ETT	-3	-6	13	10	0	12	-13	2
R-[illegible]T	-13	-10	-8	-1	-4	-6	-5	-6
R-ADC	7	4	4	12	-12	5	-4	4
R-SMT	6	1	13	8	0	14	-11	3
R-CNT	16	17	1	-7	-31	-23	-26	-4
C-ADC	-24	-26	-17	-18	-5	2	-9	-14
C-STW	4	7	13	2	15	22	-17	-3
C-PMT	-8	-12	15	3	-22	-13	0	-5
C-DTC	-6	-6	16	16	-10	-6	7	16
C-NST	2	1	14	-3	-21	-29	5	-4
R-MISS	-21	-14	-32	-25	-1[illegible]	-13	11	-14
C-MISS	-20	-12	-36	-28	-3	-8	4	-7
PRESPRY	21	18	11	13	15	-4	5	-1
SITE05	-8	8	-13	-1	-	-	-	-
SITE[illegible]5	31	20	48	32	-	-	-	-
SITE12	-21	-25	-18	-6	-44	-35	-35	-5

942

Table 41

Interactive Reading Assessment System - Spanish
Correlation Summary

SENTENCE READING

Variable	Year 1 r	Year 1 p/r	Year 2 r	Year 2 p/r	Year 3 r	Year 3 p/r	Year 4 r	Year 4 p/r
SPNCATG	35		39		36	3	36	
PREVYR	12		70		62		76	
PROGGRP	33	25	23	2	-12	-16	-11	-1
PROGYx	31	23	24	5	-11	-12	-11	6
R-QFL	3	4	-3	6	-2	-7	-16	-12
R-DGI	8	7	-13	8	1	-1	29	23
R-ETT	10	4	11	14	-5	10	-19	-14
R-NST	5	2	-12	-16	0	-1	-1	-2
R-ADC	9	7	7	20	-18	-16	-9	-6
R-SMT	8	1	2	3	-14	-14	-17	-12
R-CNT	-7	-4	-7	-13	-19	-17	-29	-23
C-ADC	-25	-23	-10	-6	-19	-10	0	2
C-STW	8	8	6	-6	3	-7	-17	-5
C-PMT	-14	-12	19	2	-19	-20	-2	-1
C-DTC	7	9	13	9	-11	-2	-19	-25
C-NST	13	10	5	-11	-4	-7	10	-1
R-MISS	-22	-16	-23	-11	-6	-3	10	-16
C-MISS	-27	-21	-33	-17	3	1	22	-1
PRESPRY	18	15	7	10	6	1	32	33
SITE05	-16	-5	-10	6	-	-	-	-
SITE35	25	9	40	25	-	-	-	-
SITE12	-16	-19	-12	-10	-19	-13	-6	4

943

Table 42

Interactive Reading Assessment System - Spanish
Correlation Summary

NARRATIVE READING COMPREHENSION

Variable	Year 1 r	Year 1 p/r	Year 2 r	Year 2 p/r	Year 3 r	Year 3 p/r	Year 4 r	Year 4 p/r
SPNCATG	24		36		39		36	
PREVYR	7		60		74		72	
PROGGRP	8	1	15	5	-12	-15	-11	-9
PROGYx	8	1	14	4	-16	-14	-10	-4
R-QFL	11	13	-1	-2	5	7	-12	-10
R-DGI	11	7	-13	-18	2	-7	14	5
R-ETT	-6	-8	9	10	-10	-9	-13	-10
R-NST	-16	-16	-17	-18	-7	-5	4	8
R-ADC	7	6	9	13	-19	-4	-9	-10
R-SMT	-8	-12	-4	-6	-9	6	-13	-10
R-CNT	16	16	-7	-8	-13	-8	-15	-5
C-ADC	-23	-23	-7	-8	-13	-4	1	7
C-STW	3	4	0	-5	6	12	-22	-16
C-PMT	2	1	25	17	-11	5	-3	-2
C-DTC	-4	-3	6	6	-17	-1	-7	-9
C-NST	-1	-3	3	-4	3	3	20	20
R-MISS	-6	-1	-14	-9	0	1	17	5
C-MISS	-11	-6	-19	-9	13	13	19	11
PRESPRY	14	11	7	7	8	-3	34	38
SITE05	-16	-8	-16	3	-	-	-	-
SITE35	11	1	23	11	-	-	-	-
SITE12	-18	-19	-13	-6	-3	12	-6	-1

Table 43

Interactive Reading Assessment System - Spanish
Correlation Summary

EXPOSITORY READING COMPREHENSION

Variable	Year 1 r	Year 1 p/r	Year 2 r	Year 2 p/r	Year 3 r	Year 3 p/r	Year 4 r	Year 4 p/r
SPNCATG	16		32		34		35	
PREVYR	6		23		66		76	
PROGGRP	20	16	16	40	-9	-15	-6	-5
PROGYx	23	20	14	1	-11	-9	-2	2
R-QFL	4	5	-3	2	3	-3	10	-7
R-DGI	-3	-3	-25	-20	3	-4	12	-1
R-ETT	4	1	9	11	-6	5	-12	-7
R-NST	-8	-9	-16	-20	-8	-6	4	8
R-ADC	7	6	5	9	-19	-12	-8	-9
R-SMT	9	5	-1	-1	-12	-7	-11	-7
R-CNT	-16	-15	-5	-7	-15	-16	-12	1
C-ADC	-43	-42	-8	-5	-10	-4	0	7
C-STW	19	19	-1	-5	5	4	-26	-23
C-PMT	12	12	21	21	-11	-10	-5	-3
C-DTC	12	12	4	4	-19	-6	-4	-4
C-NST	-6	-9	6	-2	-1	-8	24	25
R-MISS	-8	-5	-12	-4	-2	-2	16	7
C-MISS	-25	-23	-19	-7	9	3	10	4
PRESPRY	9	7	7	10	13	10	31	35
SITE05	8	14	-7	3	-	-	-	-
SITE35	17	11	24	10	-	-	-	-
SITE12	-1	-2	-4	-7	-9	-12	-9	2

945

Table 44

Interactive Reading Assessment System - Spanish
Correlation Summary

NARRATIVE LISTENING COMPREHENSION

Variable	Year 1 r	Year 1 p/r	Year 2 r	Year 2 p/r	Year 3 r	Year 3 p/r	Year 4 r	Year 4 p/r
SPNCATG	47		57		38		23	
PREVYR	23		74		64		65	
PROGGRP	36	24	45	26	9	-7	3	-1
PROGYx	37	27	43	23	0	-12	9	4
R-QFL	-2	3	-4	-1	10	5	-13	-3
R-DGI	3	-9	-5	-4	6	-6	7	2
R-ETT	3	1	-4	3	5	5	-13	-3
R-NST	0	6	4	-1	-5	-1	10	2
R-ADC	0	-5	-3	-1	-9	-7	-13	-5
R-SMT	7	1	5	1	-6	-5	-13	-3
R-CNT	2	2	6	4	-3	-8	-7	-1
C-ADC	-12	-14	-2	7	-18	-16	-7	-4
C-STW	-4	-1	-8	-4	-2	-15	-28	-14
C-PMT	-5	-11	5	3	-13	-17	-20	-5
C-DTC	-2	-1	-3	-1	-21	-17	-9	10
C-NST	6	4	4	-4	12	10	21	11
R-MISS	-24	-15	-26	-12	-7	-1	2	5
C-MISS	-31	-23	-44	-24	-4	6	-6	-3
PRESPRY	2	-4	-15	-10	1	2	-1	-6
SITE05	-31	-14	-40	-19	-	-	-	-
SITE35	39	23	51	29	-	-	-	-
SITE12	-13	-18	-18	-18	-25	-9	-32	-21

946

Table 45

Interactive Reading Assessment System - Spanish
Correlation Summary

EXPOSITORY LISTENING COMPREHENSION

Variable	Year 1 r	Year 1 p/r	Year 2 r	Year 2 p/r	Year 3 r	Year 3 p/r	Year 4 r	Year 4 p/r
SPNCATG	46		51		43		26	
PREVYR	13		60		63		58	
PROGGRP	41	30	41	72	11	-17	-1	-6
PROGYx	37	28	37	18	2	-7	12	8
R-QFL	-3	-2	-1	5	17	10	-16	-12
R-DGI	-12	-16	-11	-1	7	-1	21	19
R-ETT	6	2	-1	-1	-4	-2	-18	-14
R-NST	4	1	5	2	-6	1	5	1
R-ADC	-2	-6	-1	2	-15	-16	-12	-8
R-SMT	7	-3	5	-1	-7	-6	-16	-12
R-CNT	3	8	7	-1	-15	-24	-21	-19
C-ADC	-19	-17	-5	2	-12	1	-17	-12
C-STW	-3	-4	-12	-9	4	-12	-31	-23
C-PMT	-14	-13	3	-2	-8	-19	-21	-12
C-DTC	-6	-5	-4	3	-20	-12	-3	1
C-NST	6	1	6	-10	-2	-6	16	14
R-MISS	-24	-16	-22	-11	-10	-11	-1	-1
C-MISS	-33	-26	-41	-24	-6	1	-9	-5
PRESPRY	4	-1	-14	-10	4	5	3	2
SITE05	-26	-13	-30	-14	-	-	-	-
SITE35	40	22	45	19	-	-	-	-
SITE12	-19	-24	-17	-20	-27	-23	-31	-17

917

previous year's performance for Sentence Reading and Reading Comprehension appeared less consistent, but were generally high and tending to increase. As with English, the pattern suggests that the relative standing of individual students with respect to Spanish literacy skill becomes more set in place and predictable with increased schooling.

Nominal Reading Program Indicators

In the third regression step, the two general Spanish reading program indicators were entered: PROCGRP, the total number of years enrolled in Spanish reading (ranging from 0 to 5), and PROGYx, the dichotomous indicator of enrollment in Spanish reading in the appropriate instructional year (0 for no current enrollment and +1 for enrollment). Recall that these factors are highly correlated.

The zero-order correlations of these two variables with the Spanish IRAS deviates are somewhat complex. For oral language scales, the correlations are positive and of moderate size during the first two years (.30 to .40); the corresponding partial correlations remain of modest size (.20 to .30). The correlations with the single-word decoding scales are positive for both zero-order and partial effects, but in Year 2 only and somewhat smaller. Sentence Reading and Expository Reading Comprehension show similar relations in Year 1 only. The effects for the last two instructional years are inconsistent, generally small, and about equally often positive and negative.

Enrollment in Spanish reading programs appears to be positively related to literacy skill in the early grades, but not in the later two grade levels. As discussed in Volume 6, students are dropped from Spanish reading as Spanish literacy skills and English oral skills increase, thus leaving these programs populated by students experiencing difficulty learning to read or in acquiring oral English skills. As such, it is not surprising that length in the program is not substantially related to Spanish reading in the later grades, and may even be negatively related to some performance measures.

Instruction

The instructional indices from the current instructional year under analysis were entered in the next step of the regression, first those obtained from the observational data (RAMOS factor scores), then those derived from teacher interviews about planned instruction (Checklist factor scores). Each of these individual dimensions is discussed separately below, treating observed instruction first, then planned instruction. In general, the instructional correlates have weak effects on the Spanish IRAS deviations. This result may reflect the relatively larger number of estimated scores employed due to the infrequency of Spanish instruction, particularly in the later instructional years. The omnibus analysis used in this volume has the advantage that the sample remains fairly constant; a different picture might have emerged with regard to the effect of instructional variables if we had investigated the subset of students who received a substantial amount of Spanish instruction during a given year.

Observed instruction. The first RAMOS factor, the quality of formal language (QFL), is an index of the formal language demands placed upon the students by the instruction offered. In the first instructional year, the correlations are generally negligible with the exception of small positive partial correlations for most of the decoding based assessments (Vocabulary Decoding, Letter-sound Decoding, Letter-sound Spelling, and Narrative Reading Comprehension), averaging about .15.

Recall that the quality of decoding index (IFLTMN) had the largest loading on this factor, and with a positive weighting. As such, the finding in the first instructional year (where decoding skills are initially acquired) of positive partial correlations for the decoding based IRAS components suggests that as instruction in decoding (i.e., letter-sound correspondences) becomes more explicit, skill across the set of literacy skills having relatively large decoding components improves. Note that this result matches that found in the English data, where the quality of formal language factor showed small positive correlations with the IRAS deviates during the first instructional year.

The second RAMOS factor is an index of direct instruction delivered by an instructor aimed at groups of students rather than individuals (direct group instruction, or DGI). The only noticeable effect for this factor is a set of positive correlations with decoding skills in Year 4. The limited representativeness of this sample and the high inter-correlations between the instructional factors in this year complicate the interpretation of this isolated finding.

The third factor, engaged text time (ETT), provides a measure of reading time where students were engaged with text materials. The correlations are small and unsystematic in virtually every instructional year.

The number of students (NST), the next RAMOS factor, is an index of the number of students constituting an instructional group. This factor generally has negative effects across the four instructional years and across the nine IRAS scales. However, the partial coefficients are reduced to values near zero in most instructional years. The trend is for decreased performance with increased group size, but the relationships are weak, and generally confounded with entry performance. Engaged text (ETT) time is negatively related to this factor, further complicating interpretation.

The fifth RAMOS factor constitutes an index of the relative amount of time devoted to instruction in decoding (ADC). The correlations for this factor are negligible and inconsistent throughout. The only exception is a slight positive trend (about .15) in the partial correlations in Year 2, suggesting that in this year,

949

skill in Spanish decoding increased as more time was allocated to this skill.

The next factor provides a measure of the relative usage of secondary materials (SMT) with respect to both the quality and quantity of such usage. No consistent effects of this factor appear in the data.

The final factor, though complex, essentially gives an index of the number of management interruptions, labelled CNT for control. In the first instructional year, the correlations are generally positive at about .15 across the single-word scales, both definition and decoding, with only slightly reduced partials. The pattern is for the correlations to become more negative in subsequent years; Year 2 has no discernible pattern, and it may be that the negative values in Years 3 and 4 are dependent on peculiarities of the restricted sample. Overall interruptions requiring instructor management appear to be slightly associated with improved performance in Year 1, but problems in classroom management in Years 3 and 4 appear to be a source of poorer performance in Spanish reading.

In summary, classroom observations for Spanish reading are only weakly related to deviations in Spanish reading in this sample. Some patterns can be detected, but even in Years 1 and 2 the effects of variation in instruction on performance in Spanish are diffuse.

Planned instruction. Turning to those dimensions of instruction based on teacher plans, the first Checklist factor is a measure of the relative amount of planned time to be devoted to instruction in decoding (ADC). The effects of this factor are largely restricted to the first two instructional years (more in the former than in the latter) in which noticeably negative partial correlations are found. The pattern is not completely consistent, but it is generally pervasive and not reduced from zero-order to partial coefficients. In the first instructional year across the nine IRAS scales, the effect averages about -.25, and in the second instructional year for the decoding scales (Vocabulary Decoding, Letter-sound Decoding, Letter-sound Spelling), it averages about -.15.

Unlike the effect of the decoding quantity factor for planned instruction in English (which was associated with positive IRAS decoding deviates), here the amount of time planned for decoding instruction is negatively related to decoding skills. Recalling the discussion of the correlations between the Checklist summary indices and the set of Checklist factor scores (Volume 6), both of the student rank variables showed negative correlations with this factor, indicating that increased percentages of time were planned in decoding instruction for the lower reading groups. Thus, the finding that increased planned decoding time is associated with decreased decoding skill may simply reflect a bias in the sample for those with relatively inferior decoding skills.

The second factor, seatwork (STW), provides an index of the relative amount of time to be devoted to independent seatwork as opposed to group work. The correlations of this factor are generally inconsistent and negligible.

The third Checklist factor, PMT, is an index of the quality of the primary materials to be employed (although the factor was more complex). In the first instructional year, the partial correlations are generally negative, with the strongest values found for the decoding scales (Vocabulary Decoding, Letter-sound Decoding, and Letter-sound Spelling), averaging about -.2, and not reduced as partial coefficients. In Year 2, positive partial correlations, averaging about .2, are found for the reading comprehension scales.

The trends are for negative relationships with decoding skills and positive relationships with comprehension skills, restricted to the early instructional years. Recalling the positive correlation between this factor and the factor assessing the amount of planned decoding instruction (ADC), the relationships seem reasonable: more decoding instruction is planned for those students with the weakest decoding skills, while those students with greater decoding skills are to receive increased time devoted to comprehension activities (with both types of instruction tending to rely more heavily on the use of basals).

The next Checklist factor, decoding teacher classification (DTC), defines the relative educational training of the teacher expected to deliver decoding instruction as combined with the explicitness of such planned instruction. Little systematic effect is found for this factor. In Year 2, small positive partial correlations, averaging about .15, are found for the decoding scales (Vocabulary Decoding, Letter-sound Decoding, and Letter-sound Spelling). Thus, in the initial stages of decoding instruction, the explicitness of the decoding instruction (generally delivered by an aide) is positively related to acquired decoding instruction.

The final factor provides an index of the relative instructional group size, labelled number of students (NST). The effects are largely restricted to the last three instructional years, during which the partial correlations for both the real-word and synthetic-word decoding scales are negative, averaging about -.20. The only other scales with noticeable relationships are for the comprehension scales in the fourth instructional year, where the partial correlations for the reading comprehension scales average about .25 while those for listening comprehension average about .15. Thus, the relationships for the group size factor are generally restricted to performance in decoding, and suggest that acquisition of decoding skill is enhanced in the primary grades in this study as the relative number of students participating in an instructional group is decreased.

In summary, the correlation pattern for the five instructional factors based on teacher instructional plans, reveal the following patterns for acquisition of Spanish reading skills. First, the amount of planned decoding instruction was dependent upon the relative decoding skills of the students, with the greatest amount of such instruction planned for those students showing the weakest skills. Increased skill in decoding was associated with (a) increased explicitness in the planned decoding instruction, and (b) decreased planned group size. Comprehension skills were found to be positively related in the early instructional years to the planned increase in the use of primary materials. All of these effects are relatively slight, possibly due to artifacts mentioned earlier.

Missing instruction indicators. Following the entry of the above dimensions of instruction, observed and planned, the missing data indicators for these two sets of instructional data were entered to assess whether estimated data were playing a key role in deviate prediction. For Spanish IRAS, these indicators are proxies in part for assignment to Spanish reading; the most likely reason for missing information about instruction in Spanish was that the student had not been placed in Spanish reading. In addition, as mentioned earlier, the student might have been enrolled in a special pullout class or absent due to migration.

The largest correlations of the missing indicators are found in the first two instructional years, and occur for all of the IRAS scales except the two reading comprehension scales (where the correlations are near zero). The zero-order correlations are only partly reduced by the precursor factors, suggesting that program assignment tells only part of the story; the change is from -.24 to -.17 in Year 1, and from -.33 to -.21 in Year 2. The trend in Years 3 and 4 are also generally negative, but are smaller and less consistent. These correlations are as expected, given that missing these estimates were more typical of students not enrolled in Spanish reading programs, who thereby lacked the opportunity to advance their Spanish literacy skills.

Attendance

The average percentage of days in school attendance during instructional year x, PRESPRYx, evidenced correlations that were generally positive but small. The coefficients for reading scales in Year 1 were modest, ranging from .15 to .20 in most instances and not reduced in the partials. Reading comprehension was correlated with attendance in Year 4 (.30 to .35), but any interpretation is limited for this cohort.

Site Contrasts

In the final regression step, the three site contrast variables were entered: SITE05, SITE35, and SITE12. Given the cohort structure of the study, the first two site contrasts are only applicable within the first two instructional years.

The contrast between border and nonborder sites, SITE05, is based on the full sample of students and sites. This index was negative for

IRAS measures of oral language competence during Years 1 and 2, with zero-order values ranging from -.25 to -.40, which were reduced to about half this level when the precursors were taken into account. The pattern was inconsistent for the other IRAS scales. Given the coding of this contrast, Spanish oral language skills at exit from both first and second grade tend to be superior for students from the three border sites relative to those target students from the two nonborder sites, which is expected from the differences in the communities. More noteworthy, this effect is still observable even after kindergarten entry skills have been taken into account.

For the contrast between the nonborder sites (SITE35), positive correlations are found in both instructional years for all IRAS scales except reading comprehension. The zero-order correlations average .32 and .45 in Years 1 and 2, respectively, and the corresponding partial correlation averages .29 and .27. Even after level of Spanish language at kindergarten entry has been removed, students from the larger urban site tend to have relatively superior Spanish literacy skill than those from the less urban site. This finding is understandable, given the early exit of students from Spanish reading programs at the less urban site.

For the contrast between the border sites (SITE12), small negative partial correlations are found consistently across the first three instructional years for all of the IRAS scales except reading comprehension. The zero-order correlations average -.25, and the partial correlations -.18.

This finding indicates that, even after site differences in entry skill have been considered, the border site showing greater urban influence tends to have relatively poorer Spanish literacy skills at exit from each of the instructional years in all areas except reading comprehension. Given the stronger support for Spanish in the two border sites with relatively less urban influence, this finding is understandable.

The site contrasts have little correlation with Spanish reading comprehension, compared with their effect on other IRAS scales. This result may be due in part to the relatively low performance of students on both the narrative and expository scales. Spanish reading instruction at these sites focused on decoding at the word and sentence level, with less emphasis on comprehension of connected text, which may explain the pattern.

Summary of the Spanish Correlations

The zero-order and partial correlations between the regression predictor variables and Spanish IRAS deviates reveal the following patterns:

As was t[illegible] in the English data, Spanish language level on entry to kindergarten is associated with above average performance in each of the IRAS component literacy skills assessed during all of the early grades.

Knowledge of the names of the letters in the Spanish alphabet is weakly (though positively) related to first grade exit performance, and, unlike the situation with the English alphabet, does not serve as a general index of preschool literacy skill. However, in all subsequent instructional years, previous performance on a given IRAS scale is related to exit performance on that scale, increasingly so over grade levels. As in English, the relative standing of individual students with respect to literacy in Spanish becomes more rigid with increased schooling.

Enrollment in Spanish reading programs is positively related to the acquisition of Spanish literacy in the early grades; this association becomes negligible in the later grade levels. Given the practice of transferring the most successful students (with respect to Spanish literacy skill) in Spanish reading programs, it is understandable that length in the program is not substantially related to acquired skill in these latter grades.

Relationships for both observed and planned instructional dimensions suggest that (a) literacy skills in general are advanced by instruction that engages students in work with text materials and by limiting interruptions and (b) decoding skills are advanced by increasing the quantity and quality of decoding instruction and decreasing the number of students in an instructional group.

Attendance tends to be positively related to acquired literacy skill, but these relationships are weaker than those found in the English data.

Site contrasts, even after entry skill differences have been removed, show that Spanish literacy skill is more advanced at those border sites that provide the greatest non-school support for Spanish.

In general, the correlational structure for Spanish is less definite and more diffuse in this study than for English, partly due to the presence of a substantial number of estimated instructional values, in part because performance in Spanish reading was at a lower level than for English, and in part because program variations in Spanish were less distinctive. However, most of the noticeable patterns are sensible, and generally match or appropriately complement the findings for English.

Regression Analyses of Predictor Variables and IRAS Deviates

This section presents the results of a series of regression analyses, which assessed the combined effects of the predictors described above for each IRAS scale and each year of the study. The patterns from the zero-order and partial correlations provide some indication of the influence of the various predictors; the regressions give a picture of the joint influence of the factors. Because of the presence of multicollinearities in the data, the combined picture is not clear cut. Nonetheless, some patterns do stand out.

Multiple regression methods can serve several purposes. One product of such an analysis can be a prediction equation, in which the dependent measure is described as a weighted linear combination of the independent variables. The method can be used to determine the statistical significance of the contribution of various predictors in accounting for variation in the outcome measure.

In the present analysis, neither of these purposes is of primary importance. Rather, our intention is to use multiple regression to describe the relative contribution of different factor sets to the proportion of variance in the outcome measures, the IRAS deviates. F-ratios are also reported to give the reader an indication of the magnitude of each source relative to the residual variance, but the intention is not to highlight the statistical generalizability of the comparison.

Procedure

Separate regression runs have been carried out for each IRAS scale for both English and Spanish, using the deviation measures described above. Predictors were introduced into the regression equation stepwise in clusters:

Step 1 Language Category
Step 2 Previous Year's Achievement
Step 3 Program:
  Years in Spanish Instruction
  Spanish Instruction in Year x
Step 4 Instructional Factors from Classroom Observation (7 RAMOS factor indices)
Step 5 Instructional Factors from Teacher Interviews (5 Checklist factor indices)
Step 6 Attendance
Step 7 Site Contrasts (3 orthogonal comparisons)

Each analysis provides a substantial amount of information. In this report we will focus on only two outcomes: (a) the value of R (the total regression index) at each step and (b) the incremental F-ratio at each step. The first outcome will be further reported as R-square, which can be interpreted as the proportion of variance in the outcome measure that is accounted for by the regression equation at each step.

The data for Vocabulary Definition (English) will be used to illustrate the approach. Table 46 presents the results from the regression analyses of this measure. The initial variance and the

Table 46

Regression Analysis of Vocabulary Definition (IRAS - E):
Multiple - R and Incremental F - Ratios for Predictor Factors

Source	Year 1 R	Year 1 F	Year 2 R	Year 2 F	Year 3 R	Year 3 F	Year 4 R	Year 4 F
Language	46.0	67.0	36.6	37.9	34.5	12.2	43.8	13.4
Prev Achieve	51.0	16.9	53.9	53.6	68.9	60.3	56.8	10.6
Program	59.0	15.9	54.0	0.5	69.0	0.4	58.0	0.5
RAMOS	62.0	2.0	61.0	4.3	74.5	2.0	72.8	2.7
Checklist	67.0	5.5	68.7	8.5	78.7	2.5	80.0	2.5
Attendance	68.0	1.7	69.0	0.7	78.8	.0	83.6	2.5
Site	68.0	0.8	72.0	7.2	84.0	6.5	91.0	6.5

Incremental R Square

Source	Year 1	Year 2	Year 3	Year 4
Language	21.2	13.4	11.9	19.2
Prev Achieve	4.9	15.7	35.6	13.1
Program	8.8	0.1	0.1	1.4
RAMOS	3.6	8.1	7.9	19.4
Checklist	6.5	10.0	6.4	11.0
Site	1.4	4.6	8.6	18.8
Sum	46.2	51.8	70.6	82.8
Initial variance:	7.4	6.2	10.1	10.4
Initial df:	247	245	91	57

119

degrees of freedom for the measure are shown at the bottom of the table for each of the four years. In Year 1, the initial variance in the deviation scores was 7.4, and the 248 valid cases for the analysis provided 247 degrees of freedom.

The seven sets of predictors are listed under Sources along the left margin. In Year 1, the value of R after the first step (Language Category) was .46, with an F-ratio of 57.0 (regression values are multiplied by 100 in the tables). Previous Achievement (Alphabet Knowledge in this year) increased R to a value of .51, and the F-ratio for the incremental change in variance was 16.9. The other entries in the table can be interpreted in like fashion. The reader should keep in mind the degrees of freedom for each cluster of predictors implicit in the table of "steps" presented above: Program has 2 df, RAMOS and Checklist have 7 and 5 df respectively, and Site has 3 df in Years 1 and 2; all the other sources have 1 df each.

Lists of numbers like those in Table 46 can be a challenge to the eye. Accordingly, the information has been converted to a visual form that may be more readily interpreted. The R-value at a given step, when squared, equals the proportion of variance in the dependent measure accounted for by the regression equation at that point. At Step 1 in the example discussed above, the R-value of .46 can be squared to yield .21, which in turn means that Language Category accounts for 21 percent of the variability in the Vocabulary Definition deviations in Year 1. At Step 2, R increased to .51, which when squared yields .26, a 5 percent increment in the variance accounted for.

Figure 4 displays the increments in Percent Variance for six of the seven clusters (the graphing package available to us could accommodate only six clusters; Attendance was dropped as the factor that had least consistent and noticeable effect). The leftmost bar shows the values for Year 1 of Vocabulary Definition. The bottommost "layer" of the bar is the initial increment of 21.2 percent attributable to the Language Category cluster. Next is the increment of 5 percent from Previous Avhievement. The two Program indicators add another 8.8 percent to predictable variance; the Instructional factors contribute 10.1 percent; and the Site contrasts yield a modest 1.4 percent. The total variance accounted for by the six clusters (not including Attendance) is 46.2 percent.

Expectations Under the Null Hypothesis

In order to know how to interpret the data in the tables and figures, consider how the findings would look under the absolute null hypothesis: None of the factors has any systematic relation to the dependent measure. Each of the predictors would then be expected, on the average, to account for a proportion of the total variance equal to 1/T, where T is the total degrees of freedom. If the regression analyses had permitted completely open entry of predictors, then the findings would have been more likely to reflect random variations. Given a set of 20 predictors, one of these will account for more than

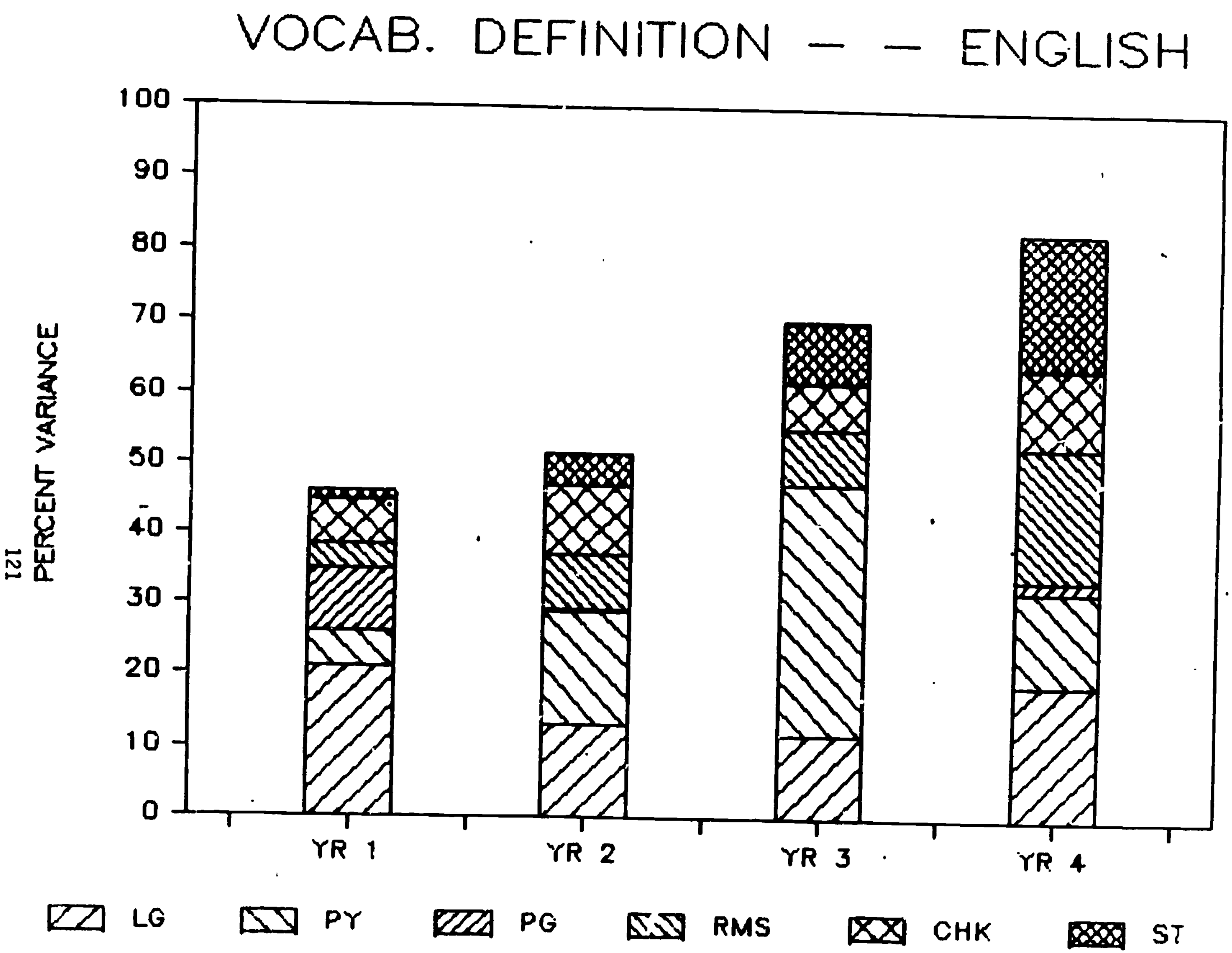


Figure 4. Distribution of explained variance in IRAS-E Vocabulary Definition.

an average amount of variance, and with open entry this factor would enter the equation first. Because a constrained order of entry was imposed in the analysis, this potential artifact was avoided.

Assuming the null hypothesis as described above, then the expected percentage of variance for each predictor degree of freedom can be determined for each of the analyses. In Years 1 and 2, there are about 250 degrees of freedom for each of the IRAS scales in both English and Spanish. Language Category and Previous Year's Achievement, each with 1 degree of freedom, should each account for .4 percent of the variance. Program, with 2 degrees of freedom, should have about .8 percent, and RAMOS and Checklist should comprise 2.8 and 2.0 percent, respectively. The total variance from all seven clusters (20 degrees of freedom) is expected to be 20/250 or 8 percent of the variance.

In Years 3 and 4, the total degrees of freedom are about 95 and 60, respectively. The expected percent of variance is 1 and 1.7, respectively, per degree of freedom, and the overall percentage from all 18 predictor degrees of freedom is 19 and 30 for Years 3 and 4, respectively (only one Site contrast applies in the last two years).

A variation in the general null hypothesis includes the assumption that an alternative hypothesis holds for one or more of the independent factors. For instance, there is good reason to believe that the precursor factors are substantially correlated with performance; that is why they are included in the analysis. In Year 1 for Vocabulary Definition, the precursors account for 26 percent of the variance, compared with the expectation of .8 percent. In this instance, it makes sense to reconditionalize the expectation for the remaining predictors. The remaining degrees of freedom are 246, the remaining variance is 74 percent of the original, and so 1/246 of 74 percent or .3 percent of the variance is expected for each of the remaining predictor degrees of freedom. The 18 remaining contrasts should add another 5.4 percent of predictable variance to the 26 percent from the precursors, for a total of 21.4. In fact, the actual contribution is 20.3, raising the total to 46.2.

The object of the analyses, then, is to assess the contribution of the various clusters of predictors to the IRAS deviations. We begin by looking for general patterns over scales and years, and then will examine more specific variations. The analyses are admittedly preliminary, and do not make complete use of the data base. Moreover, the reader needs to be mindful of certain limitations of the data base, in particular the changes in the sample cohorts after Year 2, and for Spanish the fact that estimates were included to handle missing instructional information for a substantial number of cases.

General Findings

The data from the regression analyses are displayed in Tables 46 to 54 and Figures 4 to 12 for English IRAS scales, and in Tables 55 to

63 and Figures 13 to 21 for Spanish IRAS scales. The data for Vocabulary Definition (English) are repeated for convenience.

Yearly Trends

Certain general trends appear in the tables. In both English and Spanish, there is a steady increase in the total predicted variance from Year 1 to Year 4. For English, the range in total predicted variance in Year 1 is 40 to 50 percent over IRAS scales; by Year 4, the range in total is from 80 to 90 percent. For Spanish, the corresponding totals are smaller, ranging from 25 to 50 in Year 1, and 70 to 85 in Year 4.

As noted above, we would expect under the null hypothesis an increase from Year 1 to Year 4 because of the reduction in the total degrees of freedom. The expected change would be from 8 percent to 30 percent in Years 1 and 4, respectively. The observed trends are much more substantial, and the F-ratios indicate that the omnibus null hypothesis is untenable.

The precursors are responsible for a substantial amount of the increase in variance accounted for. Recall that in Year 1, Alphabet Knowledge is used to assess previous year's achievement for all IRAS scales, while in Years 2 and subsequent the corresponding IRAS scale from the previous year served as the precursor. Language Category, the first precursor entered into the equation, remains a fairly constant predictor over years and scales; Previous Year's Achievement is responsible for most of the increase in precursor effects.

On the other hand, an interesting relation appears when the contribution of the remaining predictors is conditionalized on the percentage of variance remaining after the precursors enter the equation. As mentioned earlier for Vocabulary Definition (English), the precursors account for 26 percent of the variance in Year 1, leaving a remainder of 74 percent. The four clusters entered after the precursors comprise 18 degrees of freedom, and under the null hypothesis might be expected to account for about 7.3 percent of the remaining variance. In fact, they add another 20.3 to the total variance accounted for; looked at from another perspective, the four predictor clusters account for 20.3/74 or 27 percent of the residual variance after the precursors.

With the conditional variance as an index, for English IRAS scales the post-precursor predictors (program, instruction, attendance, and site) account for 20 to 35 percent of the residual variance in Year 1, 25 to 45 percent in Years 2 and 3, and 50 to 75 percent in Year 4. For Spanish IRAS scales, the conditional percentages range from 15 to 45 percent without any noticeable yearly trends.

To summarize these patterns, variance in both English and Spanish scales is predicted increasing well over years. A substantial amount of this trend is due to the precursor correlations, which were described in the preceding section of the volume. The remaining

Table 46

Regression Analysis of Vocabulary Definition (IRAS - E):
Multiple - R and Incremental F - Ratios for Predictor Factors

Source	Year 1 R	Year 1 F	Year 2 R	Year 2 F	Year 3 R	Year 3 F	Year 4 R	Year 4 F
Language	46.0	67.0	36.6	37.9	34.5	12.2	43.8	13.4
Prev Achieve	51.0	16.9	53.9	53.6	68.9	60.3	56.8	10.6
Program	59.0	15.9	54.0	0.5	69.0	0.4	58.0	0.5
RAMOS	62.0	2.0	61.0	4.3	74.5	2.0	72.8	2.7
Checklist	67.0	5.5	68.7	8.5	78.7	2.5	80.0	2.5
Attendance	68.0	1.7	69.0	0.7	78.8	.0	83.6	2.5
Site	68.0	0.8	72.0	7.2	84.0	6.5	91.0	6.5

Incremental R Square

	Year 1	Year 2	Year 3	Year 4
Language	21.2	13.4	11.9	19.2
Prev Achieve	4.9	15.7	35.6	13.1
Program	8.8	0.1	0.1	1.4
RAMOS	3.6	8.1	7.9	19.4
Checklist	6.5	10.0	6.4	11.0
Site	1.4	4.6	8.6	18.8
Sum	46.2	51.8	70.6	82.8
Initial variance:	7.4	6.2	10.1	10.4
Initial df:	247	245	91	57

Table 47

Regression Analysis of Narrative Listening Comprehension (IRAS - E):
Multiple - R and Incremental F - Ratios for Predictor Factors

Source	Year 1		Year 2		Year 3		Year 4	
	R	F	R	F	R	F	R	F
Language	47.00	71.10	47.00	70.72	47.00	25.87	35.80	8.24
Prev Achieve	47.70	1.18	66.50	96.42	78.80	95.05	73.60	50.07
Program	52.80	8.58	69.00	9.15	79.00	0.86	77.80	4.40
RAMOS	57.00	2.48	73.00	3.98	79.50	0.11	84.90	2.75
Checklist	61.00	3.60	76.00	4.75	82.00	2.03	89.70	3.51
Attendance	62.00	1.27	77.70	4.68	82.00	0.10	90.00	0.60
Site	63.00	2.25	80.60	9.58	87.00	6.99	94.50	6.46

Incremental R Square

Source	Year 1	Year 2	Year 3	Year 4
Language	22.09	22.09	22.09	12.82
Prev Achieve	0.66	22.13	40.00	41.35
Program	5.13	3.39	0.32	6.36
RAMOS	4.61	5.68	0.79	11.55
Checklist	4.72	4.47	4.04	8.38
Site	2.48	7.20	8.45	8.84
Sum	39.69	64.96	75.69	89.30
Initial variance:	4.51	3.83	3.90	2.32
Initial df:	247.00	246.00	92.00	57.00

Table 48

Regression Analysis of Expository Listening Comprehension (IRAS - E): Multiple - R and Incremental F - Ratios for Predictor Factors

Source	Year 1		Year 2		Year 3		Year 4	
	R	F	R	F	R	F	R	F
Language	49.00	59.21	41.00	46.38	51.00	32.39	33.90	7.27
Prev Achieve	51.60	5.43	53.00	35.13	69.80	39.66	74.50	54.93
Program	54.60	4.65	61.80	18.37	69.90	0.07	77.80	3.32
RAMOS	58.00	1.38	67.60	4.30	70.60	0.24	84.00	2.45
Checklist	61.00	2.04	72.50	6.21	77.00	3.61	88.60	2.77
Attendance	62.70	1.50	74.70	5.04	77.00	0.17	89.00	0.45
Site	64.00	1.64	78.00	8.71	83.90	7.40	93.00	5.08

Incremental R Square

	Year 1	Year 2	Year 3	Year 4
Language	24.01	16.81	26.01	11.49
Prev Achieve	2.62	11.28	22.71	44.01
Program	3.19	10.10	0.14	5.03
RAMOS	3.83	7.51	0.98	10.03
Checklist	3.57	6.86	9.45	7.94
Site	3.75	8.28	11.10	7.99
Sum	40.96	60.84	70.39	86.49
Initial variance:	4.71	5.22	4.56	3.21
Initial df:	184.00	228.00	92.00	57.00

Table 49

Regression Analysis of Vocabulary Decoding (IRAS - E):
Multiple - R and Incremental F - Ratios for Predictor Factors

Source	Year 1		Year 2		Year 3		Year 4	
	R	F	R	F	R	F	R	F
Language	38.0	42.0	40.0	47.0	47.0	26.1	42.0	12.2
Prev Achieve	55.0	57.0	73.0	193.2	82.0	122.0	82.8	89.2
Program	57.0	4.4	74.0	2.8	82.0	0.6	83.0	0.2
RAMOS	66.0	6.7	76.0	3.2	85.8	4.1	87.9	2.4
Checklist	71.0	5.9	79.6	6.8	87.0	1.5	90.6	2.3
Attendance	72.0	2.6	80.0	2.6	87.5	0.6	91.0	1.1
Site	73.0	1.6	82.0	5.6	90.0	5.0	94.9	5.9

Incremental R Square

	Year 1	Year 2	Year 3	Year 4
Language	14.4	16.0	22.1	17.6
Prev Achieve	15.8	37.3	45.2	50.9
Program	2.2	1.5	0.0	0.3
RAMOS	11.1	3.0	6.4	8.4
Checklist	6.9	5.6	2.1	4.8
Site	2.9	3.9	5.3	8.0
Sum	53.3	67.2	81.0	90.1
Initial variance:	5.8	11.8	19.1	16.6
Initial df:	248	247	92	57

Table 50

Regression Analysis of Letter Sound Decoding (IRAS - E):
Multiple - R and Incremental F - Ratios for Predictor Factors

Source	Year 1 R	Year 1 F	Year 2 R	Year 2 F	Year 3 R	Year 3 F	Year 4 R	Year 4 F
Language	31.0	26.2	27.9	20.6	43.6	21.4	44.0	13.5
Prev Achieve	41.5	22.3	57.9	94.5	71.0	58.2	87.0	131.6
Program	42.0	0.6	60.0	5.6	71.0	0.2	87.8	1.2
RAMOS	53.6	5.2	62.8	1.6	74.0	1.0	90.6	1.9
Checklist	58.0	3.7	67.6	5.4	78.0	2.4	92.0	1.5
Attendance	60.0	2.9	68.0	0.8	78.9	0.8	93.0	1.6
Site	68.0	14.4	69.0	2.6	81.6	2.7	94.0	2.1

Incremental R Square

Source	Year 1	Year 2	Year 3	Year 4
Language	9.6	7.8	19.0	19.4
Prev Achieve	7.6	25.7	31.4	56.3
Program	0.4	2.5	0.0	1.4
RAMOS	11.1	3.4	4.4	5.0
Checklist	4.9	6.3	6.1	2.6
Site	12.6	1.9	5.7	3.7
Sum	46.2	47.6	66.6	88.4
Initial variance:	4.4	5.9	5.3	4.6
Initial df:	245	245	92	57

Table 51

Regression Analysis of Letter Sound Spelling (IRAS - E): Multiple - R and Incremental F - Ratios for Predictor Factors

Source	Year 1 R	Year 1 F	Year 2 R	Year 2 F	Year 3 R	Year 3 F	Year 4 R	Year 4 F
Language	28.6	21.9	36.9	38.0	43.8	21.i	32.0	6.5
Prev Achieve	39.0	21.1	71.6	184.7	80.0	109.9	71.6	46.2
Program	42.7	4.1	72.0	2.8	80.7	1.4	72.6	0.9
RAMOS	53.6	4.9	73.7	1.4	83.0	1.6	79.6	1.9
Checklist	59.8	5.0	75.7	3.1	85.5	2.0	81.0	0.6
Attendance	61.6	2.8	76.8	3.0	86.6	1.9	83.7	1.9
Site	68.0	12.3	78.9	6.0	90.0	7.0	90.0	5.1

Incremental R Square

	Year 1	Year 2	Year 3	Year 4
Language	8.2	13.6	19.2	10.2
Prev Achieve	7.0	37.6	44.8	41.0
Program	3.0	0.6	1.1	1.4
RAMOS	10.5	2.5	3.8	1.4
Checklist	7.0	3.0	4.2	10.7
Site	10.5	4.9	7.9	2.2
				15.4
Sum	46.?	62.3	81.0	81.0
Initial variance:	194.8	409.8	536.7	519.1
Initial df:	246	241	90	57

129

Table 52

Regression Analysis of Sentence Reading (IRAS - E):
Multiple - R and Incremental F - Ratios for Predictor Factors

Source	Year 1 R	Year 1 F	Year 2 R	Year 2 F	Year 3 R	Year 3 F	Year 4 R	Year 4 F
Language	34.5	24.6	39.0	41.1	52.8	35.3	40.0	10.9
Prev Achieve	55.0	50.3	63.0	94.7	77.7	74.7	86.7	134.2
Program	58.0	4.3	70.8	22.3	77.7	0.1	87.7	2.2
RAMOS	62.5	2.2	73.8	3.0	81.9	2.4	88.0	0.3
Checklist	70.0	7.0	78.0	6.9	84.0	2.1	89.0	0.6
Attendance	71.0	1.5	79.8	5.6	84.5	0.3	89.7	0.6
Site	71.8	0.8	80.5	2.2	88.0	6.4	93.6	5.1

Incremental R Square

	Year 1	Year 2	Year 3	Year 4
Language	11.9	15.2	27.9	16.0
Prev Achieve	18.3	24.5	32.5	59.2
Program	3.4	10.4	0.0	1.7
RAMOS	5.4	4.3	6.7	0.5
Checklist	9.9	6.4	3.5	1.8
Site	2.6	4.0	6.9	8.4
Sum	51.6	64.3	77.4	87.6
Initial variance:	0.4	0.9	1.1	1.2
Initial df:	183	228	92	57

Table 53

Regression Analysis of Narrative Reading Comprehension (IRAS - E): Multiple - R and Incremental F - Ratios for Predictor Factors

Source	Year 1 R	Year 1 F	Year 2 R	Year 2 F	Year 3 R	Year 3 F	Year 4 R	Year 4 F
Language	37.00	39.68	39.00	44.26	50.00	30.76	35.70	8.20
Prev Achieve	54.00	54.48	62.00	93.44	77.00	75.75	79.00	73.86
Program	56.00	3.91	66.80	12.99	78.00	1.68	82.50	4.54
RAMOS	61.00	3.03	70.00	3.31	79.60	0.80	89.00	3.74
Checklist	65.00	4.35	73.90	4.90	81.80	1.60	91.00	1.83
Attendance	65.60	0.70	77.00	9.06	82.00	0.60	92.00	1.34
Site	66.00	0.92	78.00	4.00	89.90	14.08	95.00	5.24

Incremental R Square

Source	Year 1	Year 2	Year 3	Year 4
Language	13.69	15.21	25.00	12.74
Prev Achieve	15.47	23.23	34.29	49.67
Program	2.20	6.18	1.55	5.65
RAMOS	5.85	4.38	2.52	11.15
Checklist	5.04	5.61	3.55	3.60
Site	1.31	6.23	13.91	7.44
Sum	43.56	60.84	80.82	90.25
Initial variance:	1.96	5.06	7.03	5.40
Initial df:	248.00	246.00	92.00	57.00

Table 54

Regression Analysis of Expository Reading Comprehension (IRAS - E):
Multiple - R and Incremental F - Ratios for Predictor Factors

Source	Year 1 R	Year 1 F	Year 2 R	Year 2 F	Year 3 R	Year 3 F	Year 4 R	Year 4 F
Language	28.8	16.4	36.8	35.7	53.0	36.0	36.8	8.8
Prev Achieve	49.0	38.2	53.6	48.2	70.0	37.7	75.9	57.2
Program	50.0	1.1	62.0	18.0	72.0	2.2	81.6	7.2
RAMOS	52.7	0.9	66.9	3.5	75.0	1.2	85.0	1.5
Checklist	62.8	6.4	73.6	8.7	81.6	4.6	87.0	1.0
Attendance	63.0	0.5	75.8	5.5	81.7	0.2	87.8	0.7
Site	64.0	1.0	77.5	4.3	89.0	12.4	92.0	4.5

Incremental R Square

	Year 1	Year 2	Year 3	Year 4
Language	8.3	13.5	28.1	13.5
Prev Achieve	15.7	15.2	20.9	44.1
Program	1.0	9.7	2.8	9.0
RAMOS	2.8	6.3	4.4	5.7
Checklist	11.7	9.4	10.3	3.4
Site	1.5	5.9	12.6	9.0
Sum	41.0	60.1	79.2	84.6
Initial variance:	1.6	5.4	8.0	6.7
Initial df:	181	228	92	57

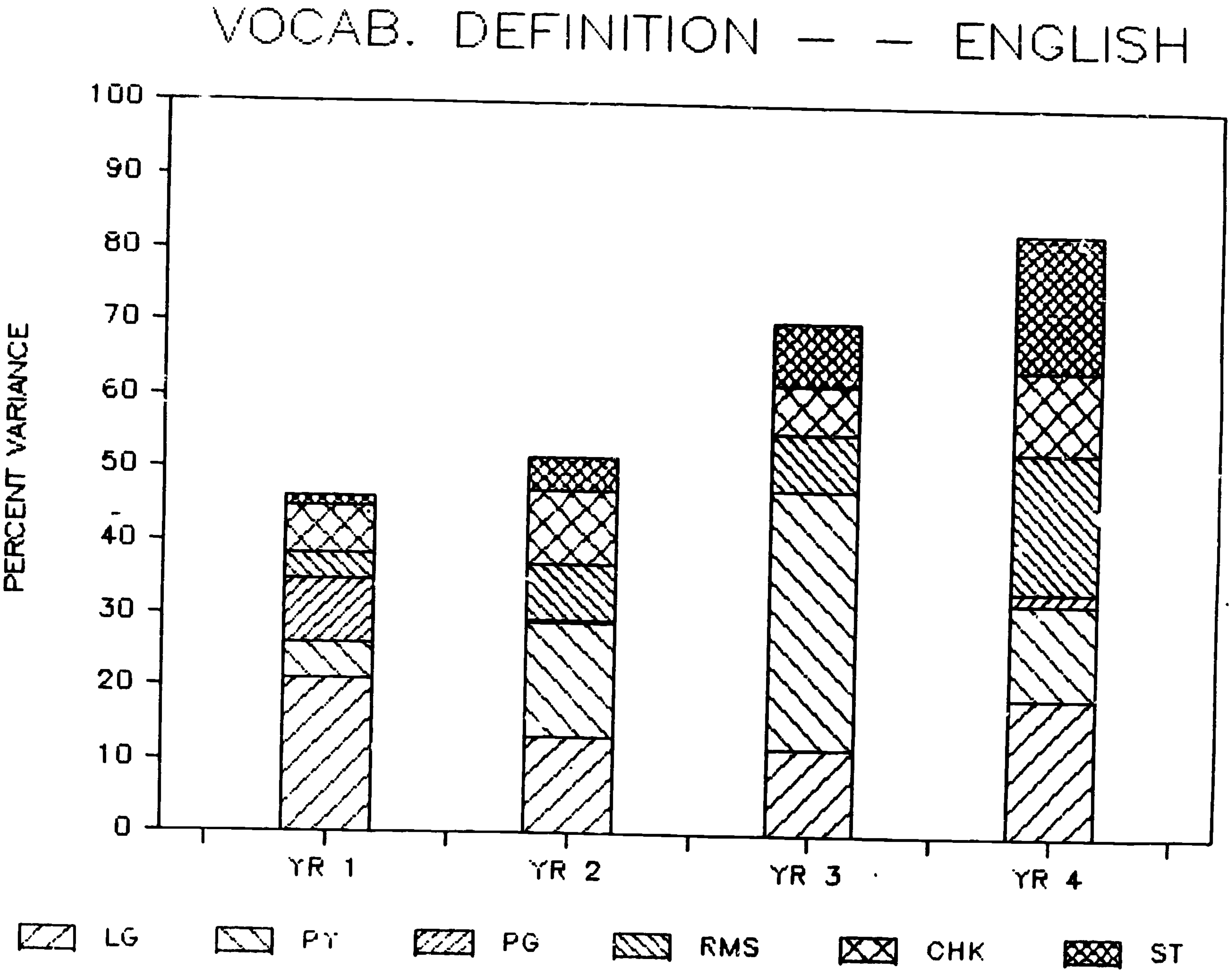
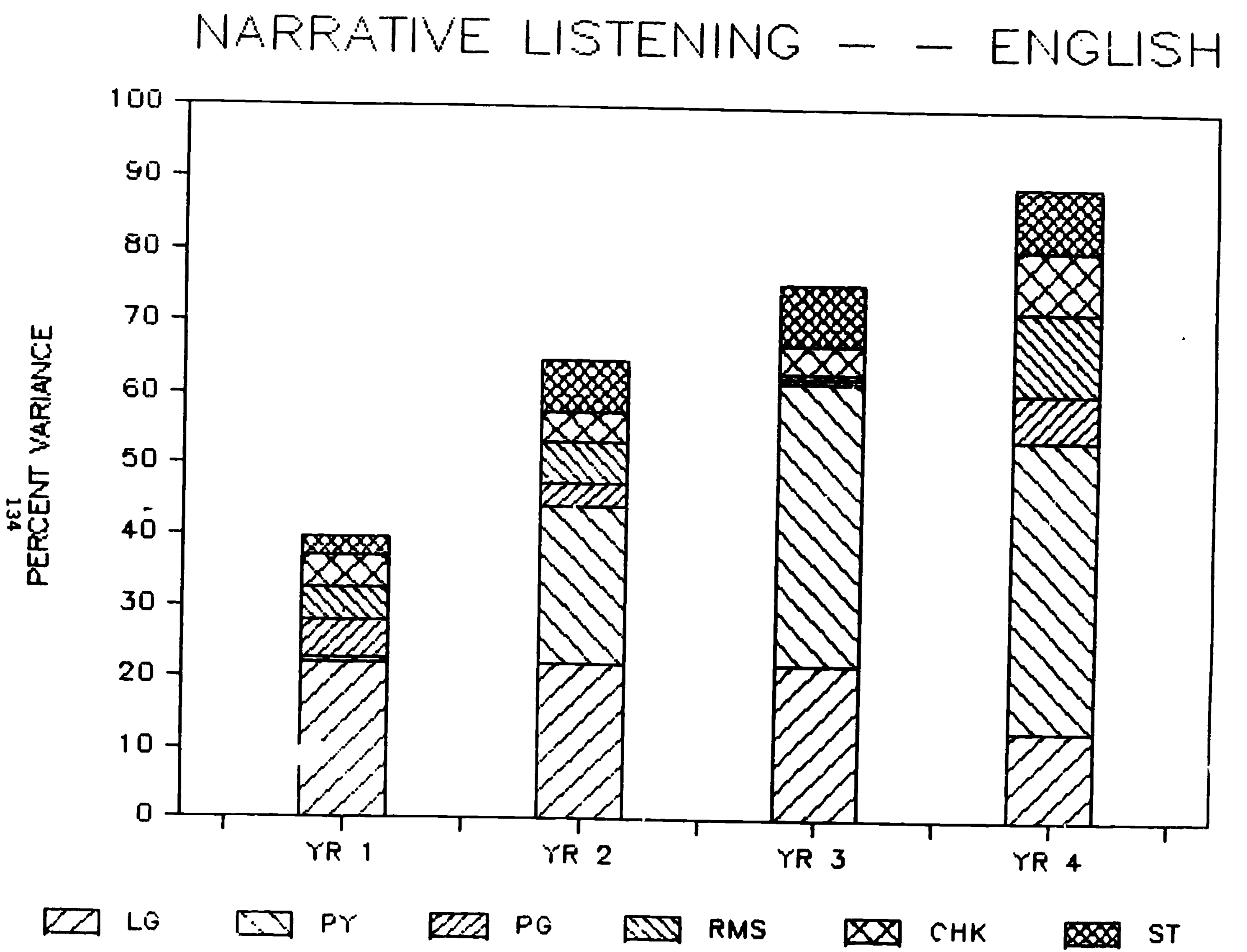


Figure 4. Distribution of explained variance in IRAS-E Vocabulary Definition.



134

Figure 5. Distribution of explained variance in IRAS-E Narrative Listening Comprehension.

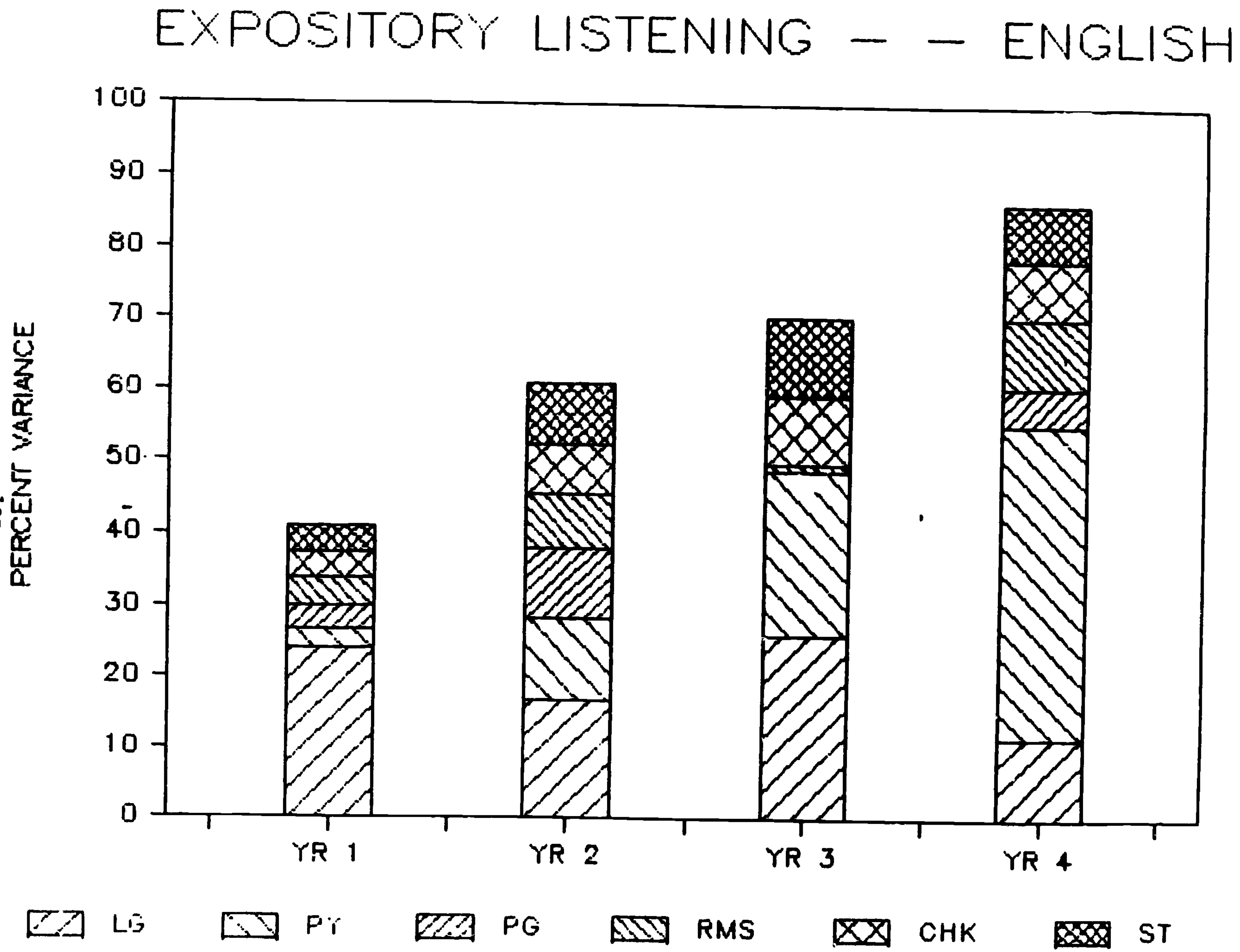


Figure 6. Distribution of explained variance in IRAS-E Expository Listening Comprehension.

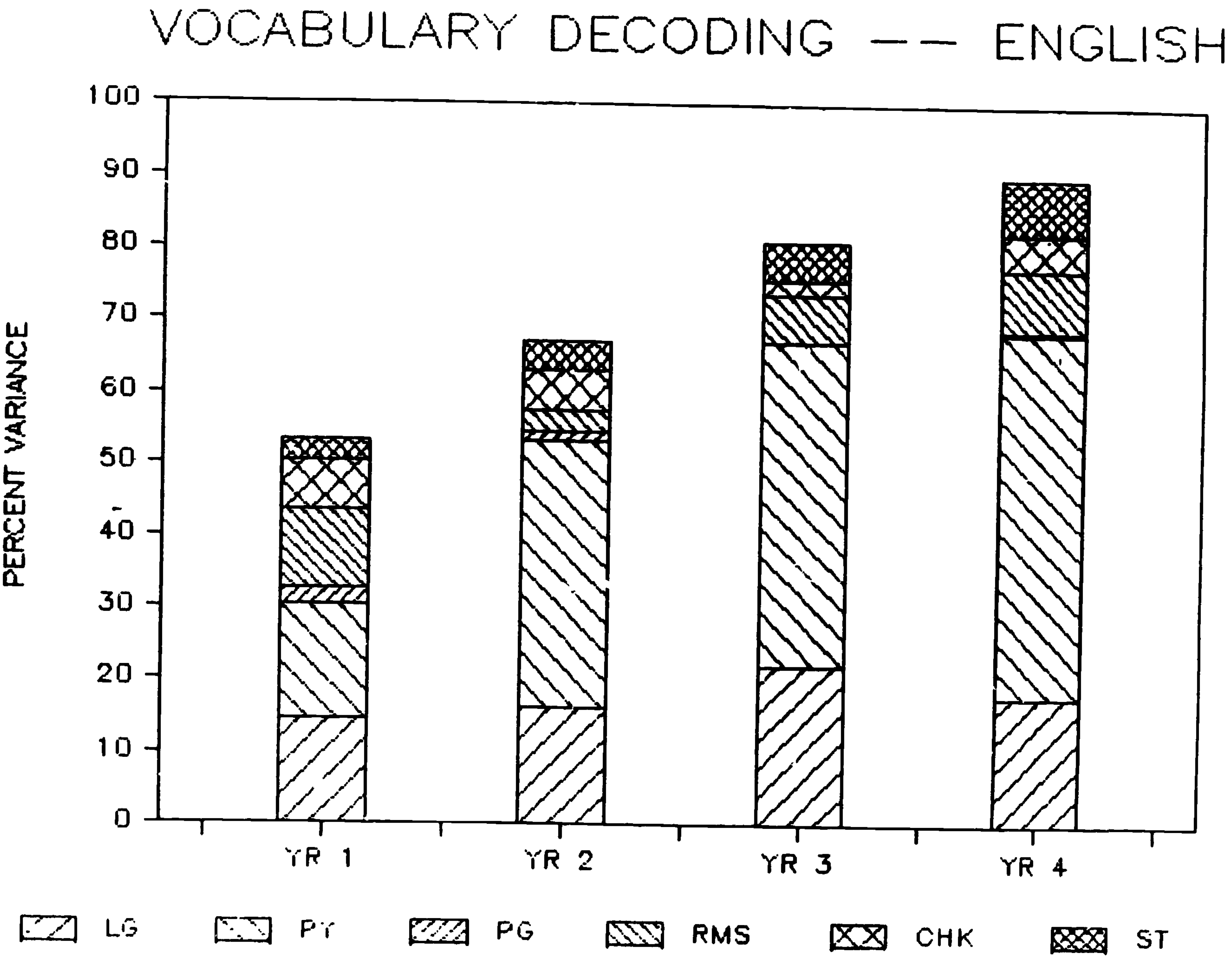


Figure 7. Distribution of explained variance in IRAS-E Vocabulary Decoding.

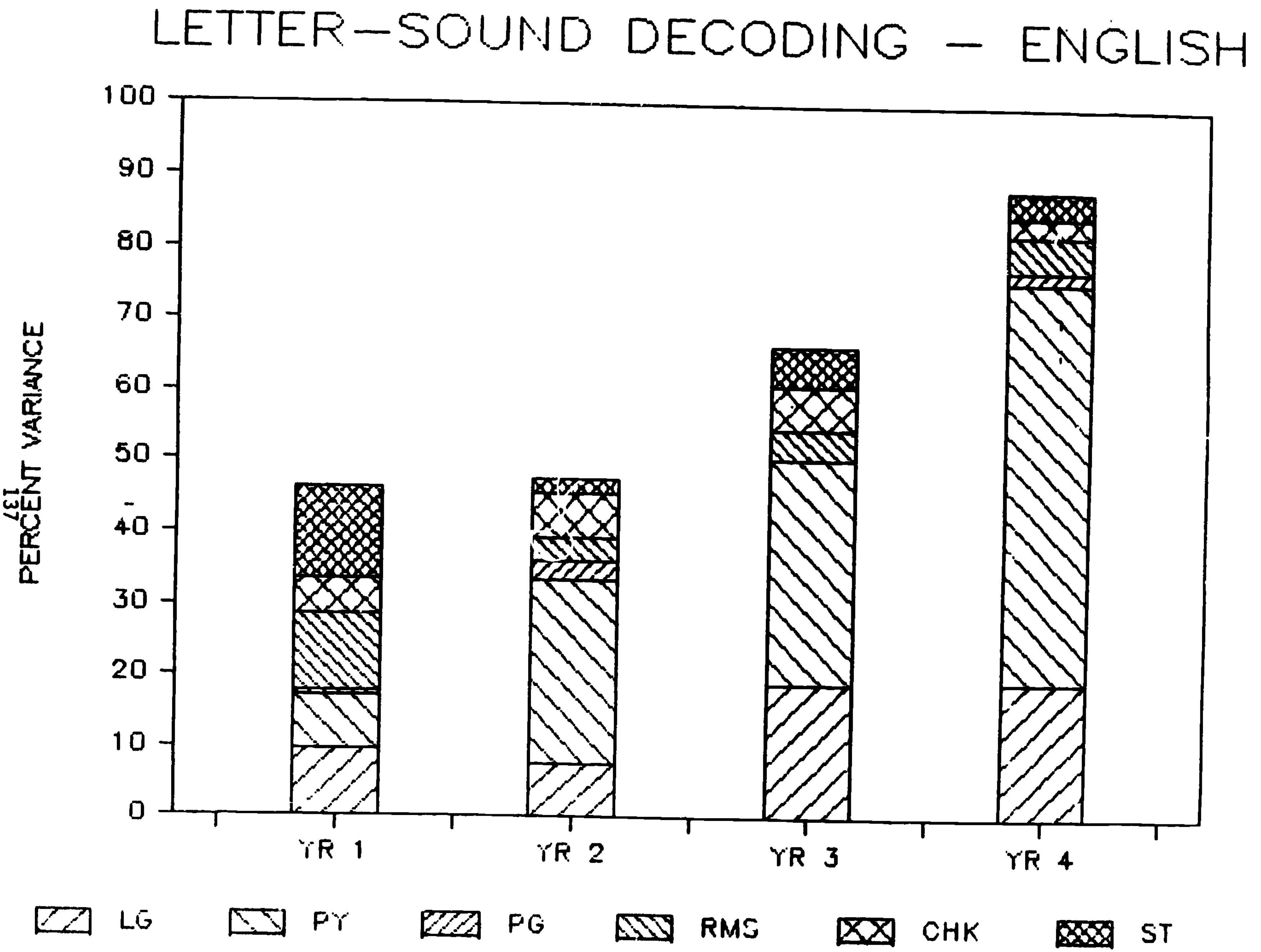


Figure 8. Distribution of explained variance in IRAS-E Letter-sound Decoding.

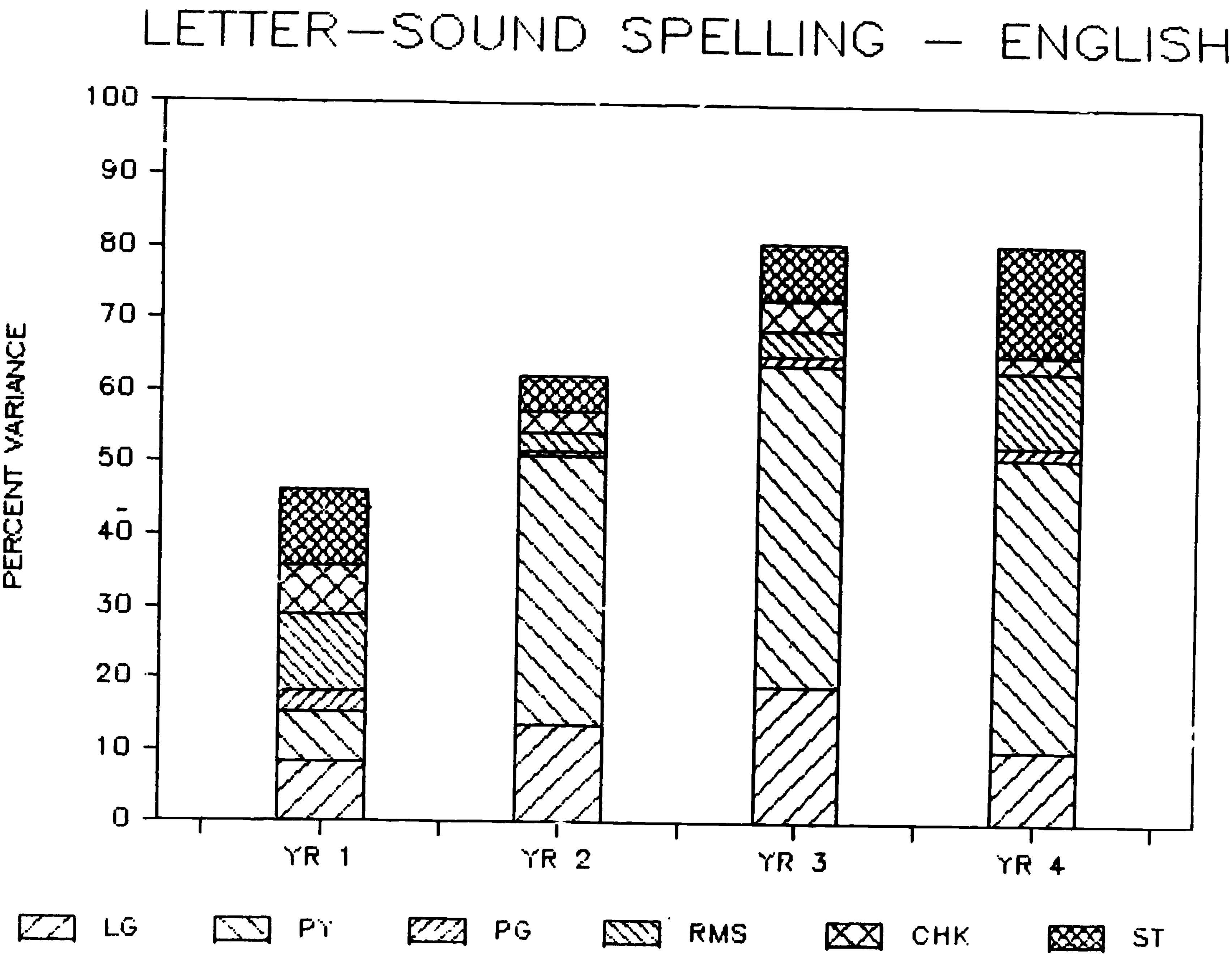


Figure 9. Distribution of explained variance in IRAS-E Letter-sound Spelling.

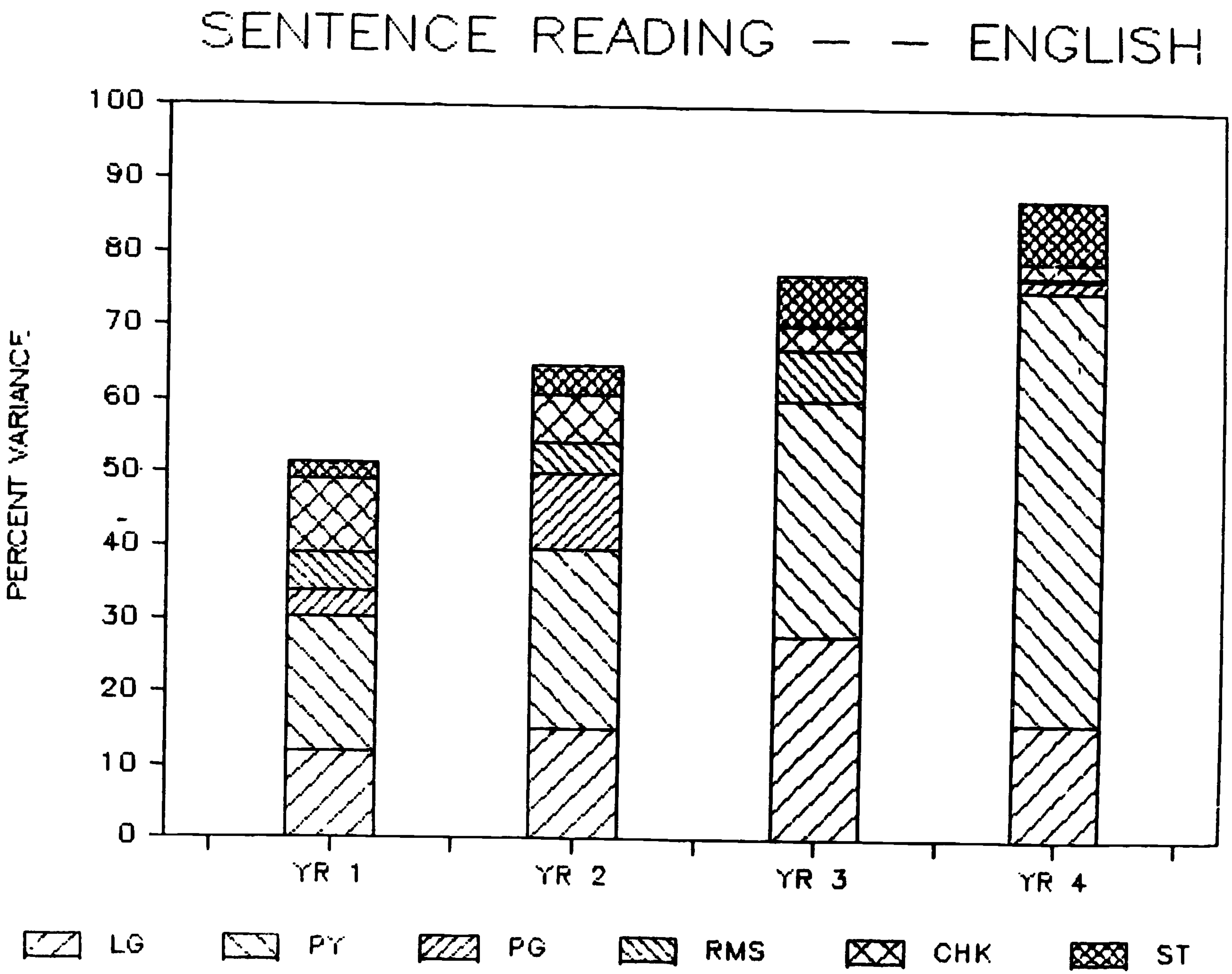


Figure 10. Distribution of explained variance in IRAS-E Sentence Reading.

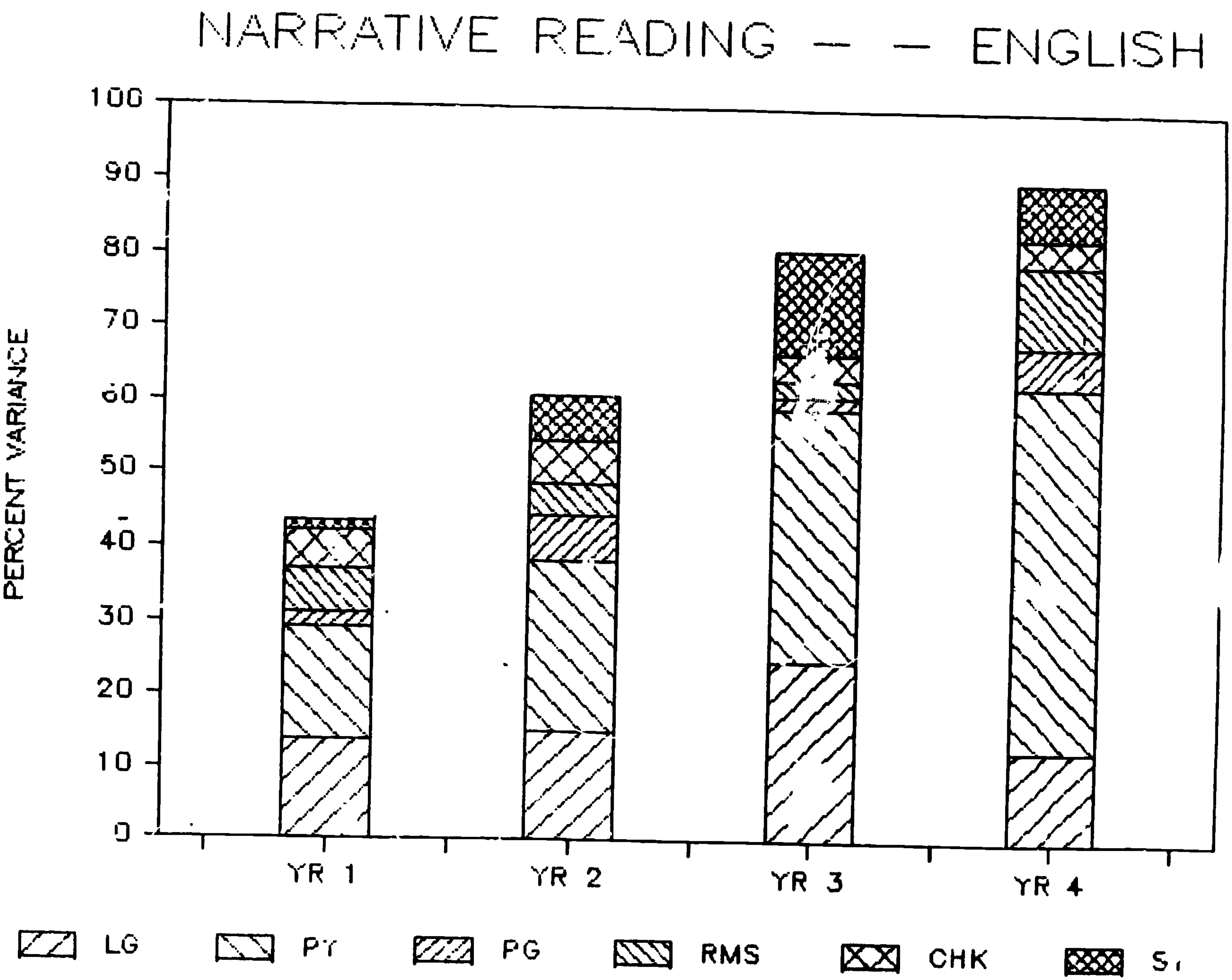


Figure 11. Distribution of explained variance in IRAS-E Narrative Reading Comprehension.

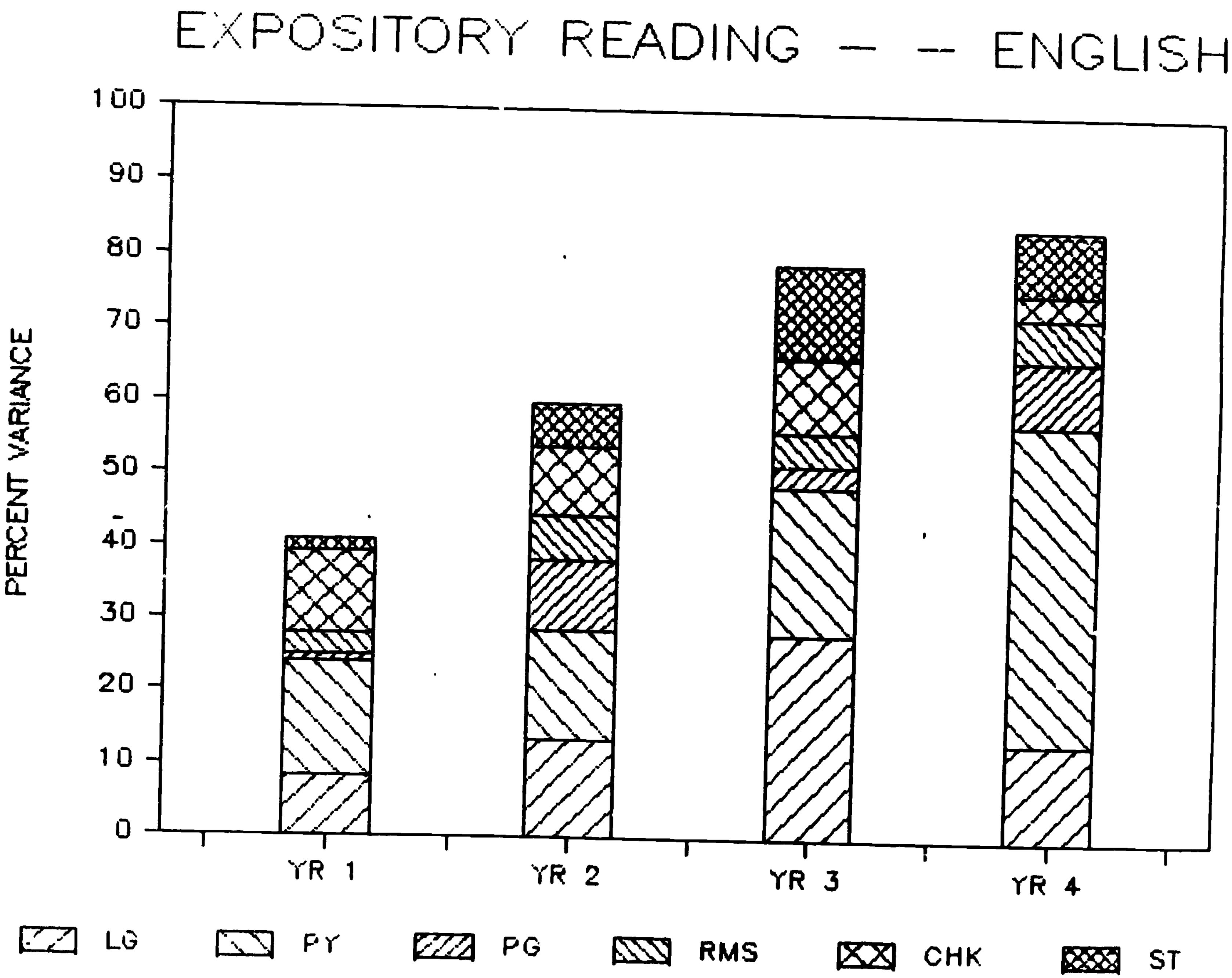


Figure 12. Distribution of explained variance in IRAS-E Expository Reading Comprehension.

997

Table 55

Regression Analysis of Vocabulary Definition (IRAS - S): Multiple - R and Incremental F - Ratios for Predictor Factors

Source	Year 1 R	Year 1 F	Year 2 R	Year 2 F	Year 3 R	Year 3 F	Year 4 R	Year 4 F
Language	45.0	58.7	42.8	54.9	24.0	5.9	36.0	8.2
Prev Achieve	46.0	3.1	63.6	91.1	61.9	48.5	72.0	58.9
Program	52.0	9.6	65.0	4.3	62.0	0.4	74.0	1.7
RAMOS	56.9	2.3	67.0	1.6	66.8	1.2	75.0	0.3
Checklist	59.0	1.9	68.8	2.0	70.0	1.4	76.0	0.3
Attendance	60.0	1.0	70.0	2.9	71.0	0.9	80.0	2.4
Site	65.8	8.7	71.9	3.6	74.0	2.1	82.8	1.6

Incremental R Square

	Year 1	Year 2	Year 3	Year 4
Language	20.3	18.3	5.8	13.0
Prev Achievr	0.9	22.1	32.6	38.9
Program	5.9	1.8	0.1	2.9
RAMOS	5.3	2.6	6.2	1.5
Checklist	2.4	2.4	4.4	1.5
Site	8.5	4.4	5.8	10.8
Sum	43.3	51.7	54.8	68.6
Initial variance:	12.1	14.9	11.3	11.4
Initial df:	231	246	94	56

Table 56

Regression Analysis of Narrative Listening Comprehension (IRAS - S):
Multiple - R and Incremental F - Ratios for Predictor Factors

Source	Year 1 R	Year 1 F	Year 2 R	Year 2 F	Year 3 R	Year 3 F	Year 4 R	Year 4 F
Language	46.6	68.2	56.8	117.1	38.0	15.9	23.0	3.1
Prev Achieve	48.0	5.3	77.7	173.7	67.0	51.9	64.6	34.6
Program	53.6	9.3	79.0	8.7	68.0	0.9	64.9	0.2
RAMOS	54.9	0.7	79.6	0.4	68.7	0.2	65.0	0.0
Checklist	57.0	1.6	80.0	0.8	73.0	2.2	66.9	0.5
Attendance	57.9	1.2	80.0	0.9	73.5	0.2	67.7	0.3
Site	62.5	6.9	81.8	5.3	79.5	5.3	83.8	9.9

Incremental R Square

Source	Year 1	Year 2	Year 3	Year 4
Language	21.7	32.3	14.4	5.3
Prev Achieve	1.3	28.1	30.5	36.4
Program	5.7	2.0	1.4	0.4
RAMOS	1.4	1.0	1.0	0.1
Checklist	2.3	0.6	6.1	2.5
Site	6.6	2.9	9.9	25.5
Sum	39.1	66.9	63.2	70.2
Initial variance:	3.8	5.0	2.8	2.1
Initial df:	247	246	94	57

Table 57

Regression Analysis of Expository Listening Comprehension (IRAS - S):
Multiple - R and Incremental F - Ratios for Predictor Factors

Source	Year 1 R	Year 1 F	Year 2 R	Year 2 F	Year 3 R	Year 3 F	Year 4 R	Year 4 F
Language	46.0	50.2	51.0	81.3	42.5	20.6	25.5	3.9
Prev Achieve	46.6	0.7	66.5	73.9	69.0	52.5	58.0	22.7
Program	54.0	9.7	69.0	7.1	69.8	16.8	60.8	1.3
RAMOS	57.0	1.4	69.0	0.2	72.0	0.8	62.8	0.3
Checklist	59.8	1.4	70.5	1.4	74.0	1.0	69.0	1.6
Attendance	60.9	1.3	71.8	2.7	74.8	0.6	70.5	0.5
Site	64.8	4.6	73.5	3.5	81.0	6.4	82.0	6.6

Incremental R Square

Source	Year 1	Year 2	Year 3	Year 4
Language	21.2	26.0	18.1	6.5
Prev Achieve	0.6	18.2	29.5	27.1
Program	7.4	3.4	1.1	3.3
RAMOS	3.3	0.0	3.1	2.5
Checklist	3.3	2.1	3.1	8.2
Site	6.2	4.3	10.9	19.6
Sum	42.0	54.0	65.6	67.2
Initial variance:	3.1	5.1	3.8	2.7
Initial df:	185	229	94	57

Table 58

Regression Analysis of Vocabulary Decoding (IRAS - S):
Multiple - R and Incremental F -- Ratios for Predictor Factors

Source	Year 1 R	Year 1 F	Year 2 R	Year 2 F	Year 3 R	Year 3 F	Year 4 R	Year 4 F
Language	31.5	27.3	36.0	37.1	32.0	10.7	30.8	5.9
Prev Achieve	31.5	0.1	75.0	242.4	83.8	185.9	82.0	100.8
Program	33.9	2.1	75.8	3.4	83.8	0.0	83.9	2.1
RAMOS	38.0	1.3	77.5	2.2	86.8	2.5	85.7	0.9
Checklist	54.0	9.6	80.0	5.8	87.0	0.6	86.0	0.2
Attendance	56.6	3.1	82.0	6.7	87.5	0.4	90.0	6.1
Site	59.6	4.1	83.8	6.9	89.0	2.8	93.0	5.5

Incremental R Square

	Year 1	Year 2	Year 3	Year 4
Language	9.9	13.0	10.2	9.5
Prev Achieve	0.0	43.3	60.0	57.8
Program	1.6	1.2	0.0	3.2
RAMOS	2.9	2.6	5.1	3.1
Checklist	14.7	3.9	0.3	0.5
Site	6.4	6.2	3.5	12.5
Sum	35.5	70.2	79.2	86.5
Initial variance:	11.6	25.7	30.3	28.7
Initial [illegible]:	248	247	94	57

Table 59

Regression Analysis of Letter Sound Decoding (IRAS - S):
Multiple - R and Incremental F - Ratios for Predictor Factors

Source	Year 1 R	Year 1 F	Year 2 R	Year 2 F	Year 3 R	Year 3 F	Year 4 R	Year 4 F
Language	35.7	35.5	36.0	37.9	27.0	7.3	18.0	1.9
Prev Achieve	35.7	.0	69.5	167.2	71.9	84.1	80.8	99.1
Program	39.0	3.6	70.7	3.7	72.0	0.3	81.0	2.9
RAMOS	45.6	2.3	73.5	3.0	75.8	1.5	82.7	0.6
Checklist	54.0	5.7	76.0	4.1	79.8	2.7	83.0	0.2
Attendance	56.8	3.1	77.0	3.4	81.7	2.3	84.0	1.3
Site	60.5	5.0	78.9	5.1	85.0	4.1	86.8	2.0

Incremental R Square

	Year 1	Year 2	Year 3	Year 4
Language	12.7	13.0	7.3	3.2
Prev Achieve	0.0	35.3	44.4	62.0
Program	2.5	1.7	0.1	0.3
RAMOS	5.6	4.0	5.6	2.8
Checklist	8.4	3.7	6.2	0.5
Site	7.4	4.5	8.6	6.5
Sum	36.6	62.3	72.3	75.3
Initial variance:	2.6	4.1	3.6	3.1
Initial df:	243	248	93	57

Table 60

Regression Analysis of Letter Sound Spelling (IRAS - S):
Multiple - R and Incremental F - Ratios for Predictor Factors

Source	Year 1 R	Year 1 F	Year 2 R	Year 2 F	Year 3 R	Year 3 F	Year 4 R	Year 4 F
Language	32.6	29.7	41.0	49.7	27.9	7.7	34.9	7.5
Prev Achieve	34.5	3.5	75.0	220.0	8[illegible].5	145.8	81.0	84.7
Program	34.9	0.4	77.0	9.1	81.0	1.[illegible]	81.3	0.2
RAMOS	41.9	2.2	78.0	1.3	83.9	1.8	81.6	.0
Checklist	50.5	5.0	79.5	2.5	83.0	1.4	82.5	0.4
Attendance	54.0	4.0	81.5	7.3	86.0	1.5	84.0	1.3
Site	57.7	4.7	83.5	7.7	89.0	4.9	65.5	1.2

Incremental R Square

	Year 1	Year 2	Year 3	Year 4
Language	10.6	16.8	7.8	12.2
Prev Achieve	1.3	39.4	57.0	53.4
Program	0.3	3.0	0.8	0.8
RAMOS	5.4	1.6	4.8	0.2
Checklist	7.9	2.4	1.9	1.5
Site	7.8	6.5	7.0	5.0
Sum	33.3	69.7	74.2	73.1
Initial variance:	351.9	597.9	586.7	578.1
Initial df:	249	243	92	55

Table 61

Regression Analysis of Sentence Reading (IRAS - S):
Multiple - R and Incremental F - Ratios for Predictor Factors

Source	Year 1 R	Year 1 F	Year 2 R	Year 2 F	Year 3 R	Year 3 F	Year 4 R	Year 4 F
Language	35.0	25.9	38.5	39.8	36.0	13.8	35.0	8.1
Prev Achieve	35.7	1.0	72.0	179.7	65.0	47.2	76.7	62.2
Program	42.5	6.1	72.5	0.8	66.5	1.4	77.0	0.3
RAMOS	44.8	0.6	74.7	2.4	69.0	0.9	78.5	0.4
Checklist	50.6	2.6	75.6	1.3	71.0	0.9	82.0	1.7
Attendance	52.6	1.6	79.0	10.5	72.8	1.1	85.6	3.4
Site	55.0	2.2	81.0	6.4	79.8	6.3	87.0	1.2

Incremental R Square

	Year 1	Year 2	Year 3	Year 4
Language	12.3	14.8	13.0	12.3
Prev Achieve	0.5	37.0	29.3	46.6
Program	5.3	0.7	2.0	0.5
RAMOS	2.0	3.2	3.4	2.3
Checklist	5.5	1.4	2.8	5.6
Site	4.6	8.5	13.3	8.5
Sum	30.3	65.6	63.7	75.7
Initial variance:	0.2	0.8	1.4	1.4
Initial df:	185	229	94	57

Table 62

Regression Analysis of Narrative Reading Comprehension (IRAS - S):
Multiple - R and Incremental F - Ratios for Predictor Factors

Source	Year 1 R	Year 1 F	Year 2 R	Year 2 F	Year 3 R	Year 3 F	Year 4 R	Year 4 F
Language	23.7	14.9	35.5	35.7	38.8	16.0	36.0	8.5
Prev Achieve	23.8	0.1	63.9	118.6	75.6	90.8	72.5	45.9
Program	23.9	0.1	64.0	0.3	76.0	1.0	72.9	0.3
RAMOS	35.0	2.7	69.5	4.8	77.5	0.6	73.0	0.1
Checklist	43.5	3.7	72.0	4.0	78.0	0.4	77.7	1.6
Attendance	44.7	1.1	73.0	1.2	78.0	0.2	82.0	3.1
Site	49.0	4.3	75.0	5.1	83.6	5.5	85.5	2.6

Incremental R Square

Source	Year 1	Year 2	Year 3	Year 4
Language	5.6	12.6	15.1	13.0
Prev Achieve	.0	28.2	42.1	39.6
Program	.0	0.1	0.6	0.6
RAMOS	6.5	7.3	2.3	0.1
Checklist	6.7	3.5	0.8	7.1
Site	5.1	4.4	9.0	12.7
Sum	24.0	56.3	69.9	73.1
Initial variance:	0.4	2.0	5.2	7.7
Initial df:	249	248	94	

Table 63

Regression Analysis of Expository Reading Comprehension (IRAS - S): Multiple - R and Incremental F - Ratios for Predictor Factors

Source	Year 1 R	Year 1 F	Year 2 R	Year 2 F	Year 3 R	Year 3 F	Year 4 R	Year 4 F
Language	15.6	4.6	31.7	25.6	33.5	11.8	35.0	7.9
Prev Achieve	15.9	0.2	36.8	9.2	68.0	61.6	76.0	61.0
Program	25.6	3.9	37.0	0.6	69.6	1.4	76.7	0.3
RAMOS	41.0	3.1	48.5	3.9	71.0	0.6	76.9	0.1
Checklist	62.0	12.2	54.6	3.8	72.0	0.5	81.0	1.9
Attendance	67.9	7.9	56.6	2.3	73.0	0.5	84.0	2.4
Site	69.5	2.4	64.0	10.7	82.0	9.0	86.0	1.8

Incremental R Square

Source	Year 1	Year 2	Year 3	Year 4
Language	2.4	10.0	11.2	12.3
Prev Achieve	0.1	3.5	35.0	45.5
Program	4.0	0.1	2.2	1.1
RAMOS	10.3	9.8	2.0	0.3
Checklist	21.6	6.3	1.4	6.5
Site	9.9	11.1	15.4	8.4
Sum	48.3	41.0	67.2	74.0
Initial variance:	0.1	1.6	4.9	8.5
Initial df:	185	229	94	57

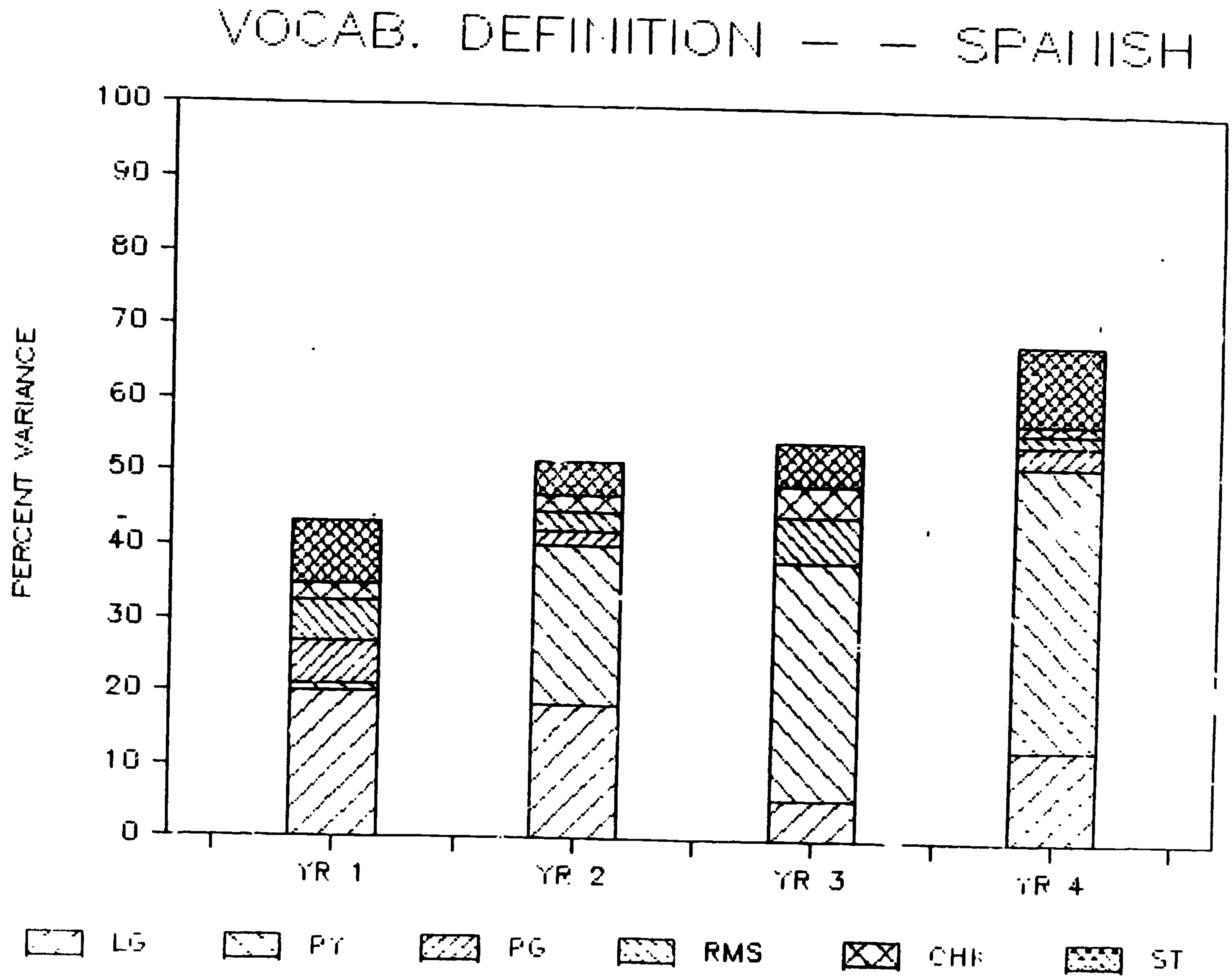


Figure 13. Distribution of explained variance in IRAS-S Vocabulary Definition.

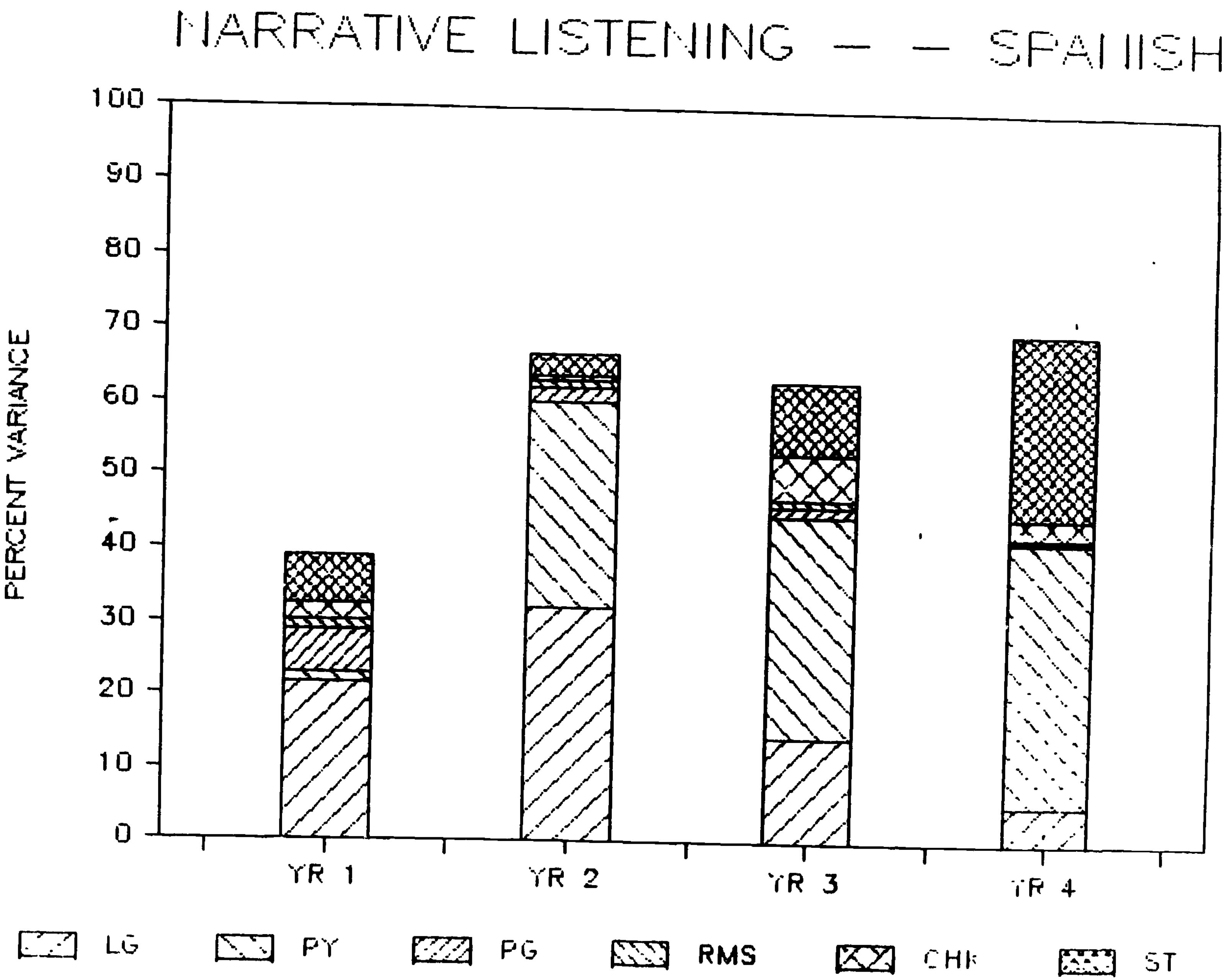


Figure 14. Distribution of explained variance in IRAS-S Narrative Listening Comprehension.

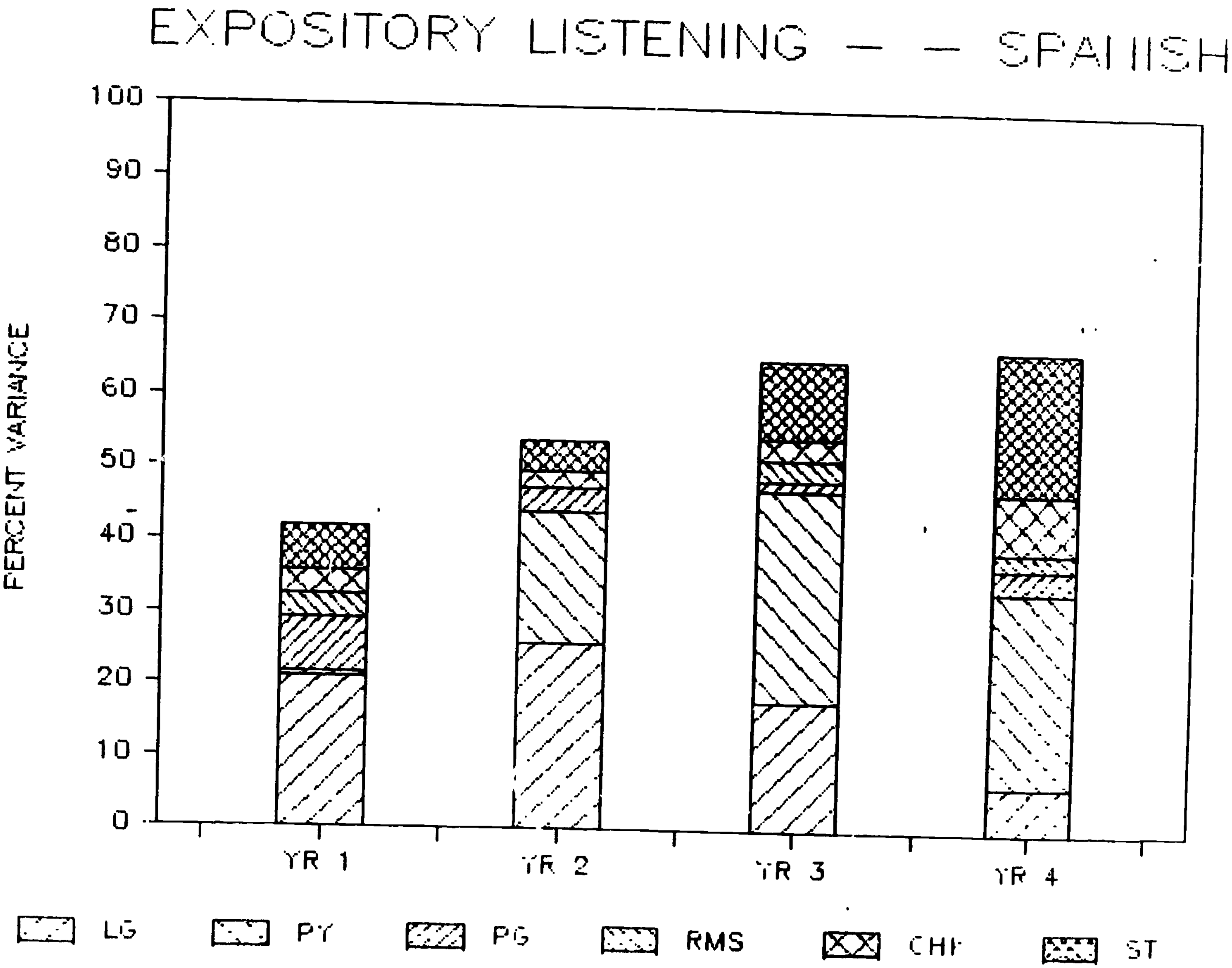


Figure 15. Distribution of explained variance in IRAS-S Expository Listening Comprehension.

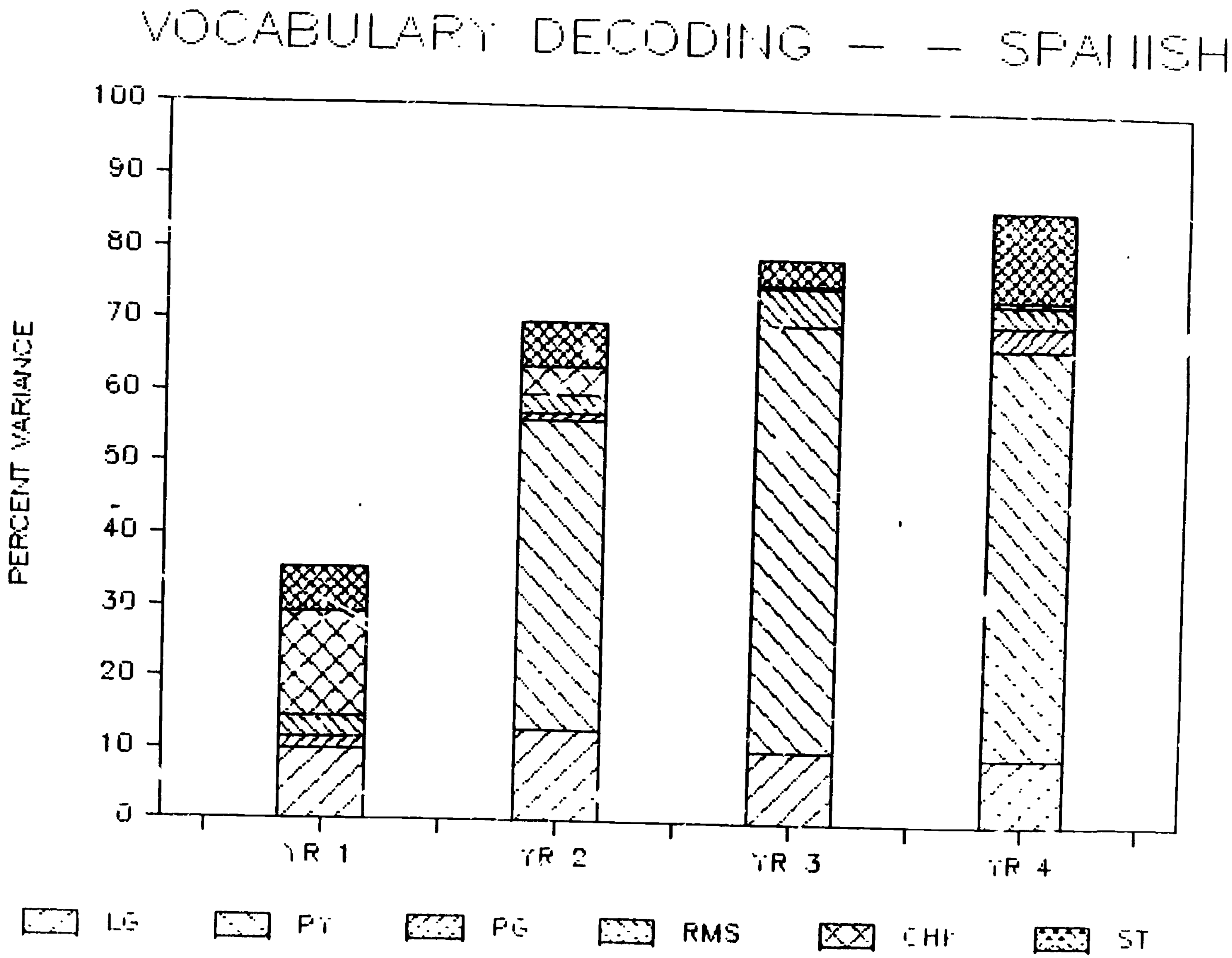


Figure 16. Distribution of explained variance in IRAS-S Vocabulary Decoding.

154

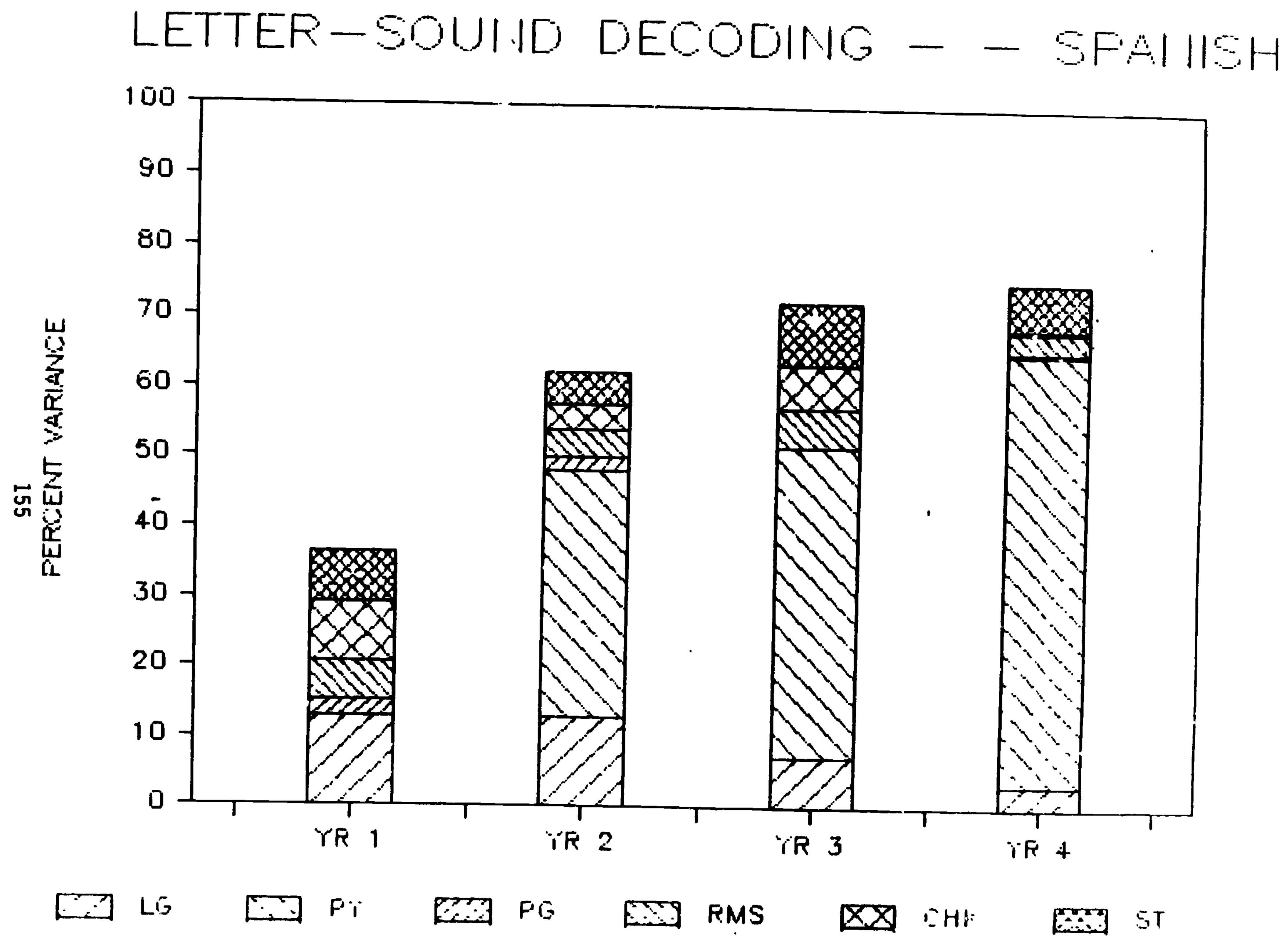


Figure 17. Distribution of explained variance in IRAS-S Letter-sound Decoding.

156

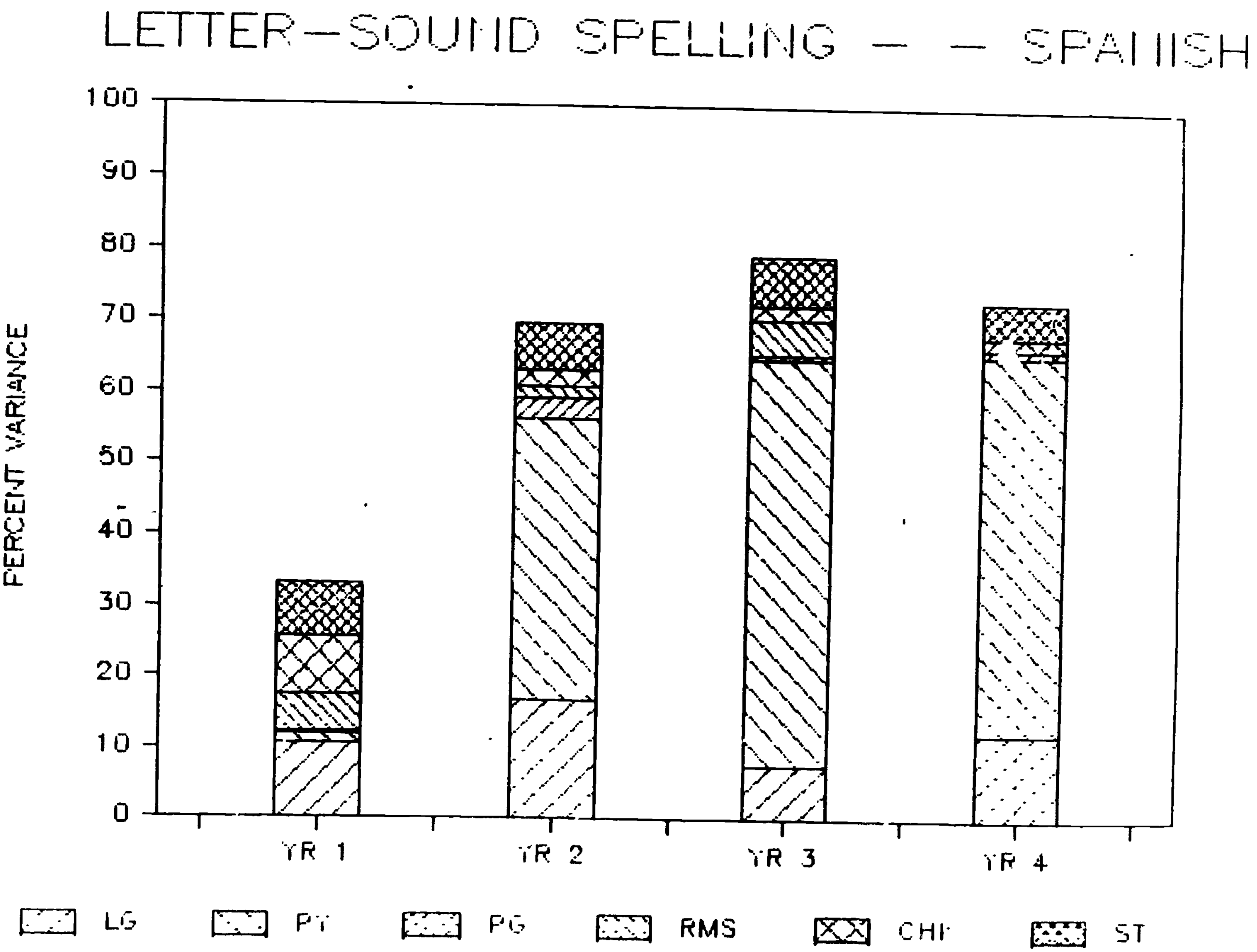


Figure 18. Distribution of explained variance in IRAS-S Letter-sound Spelling.

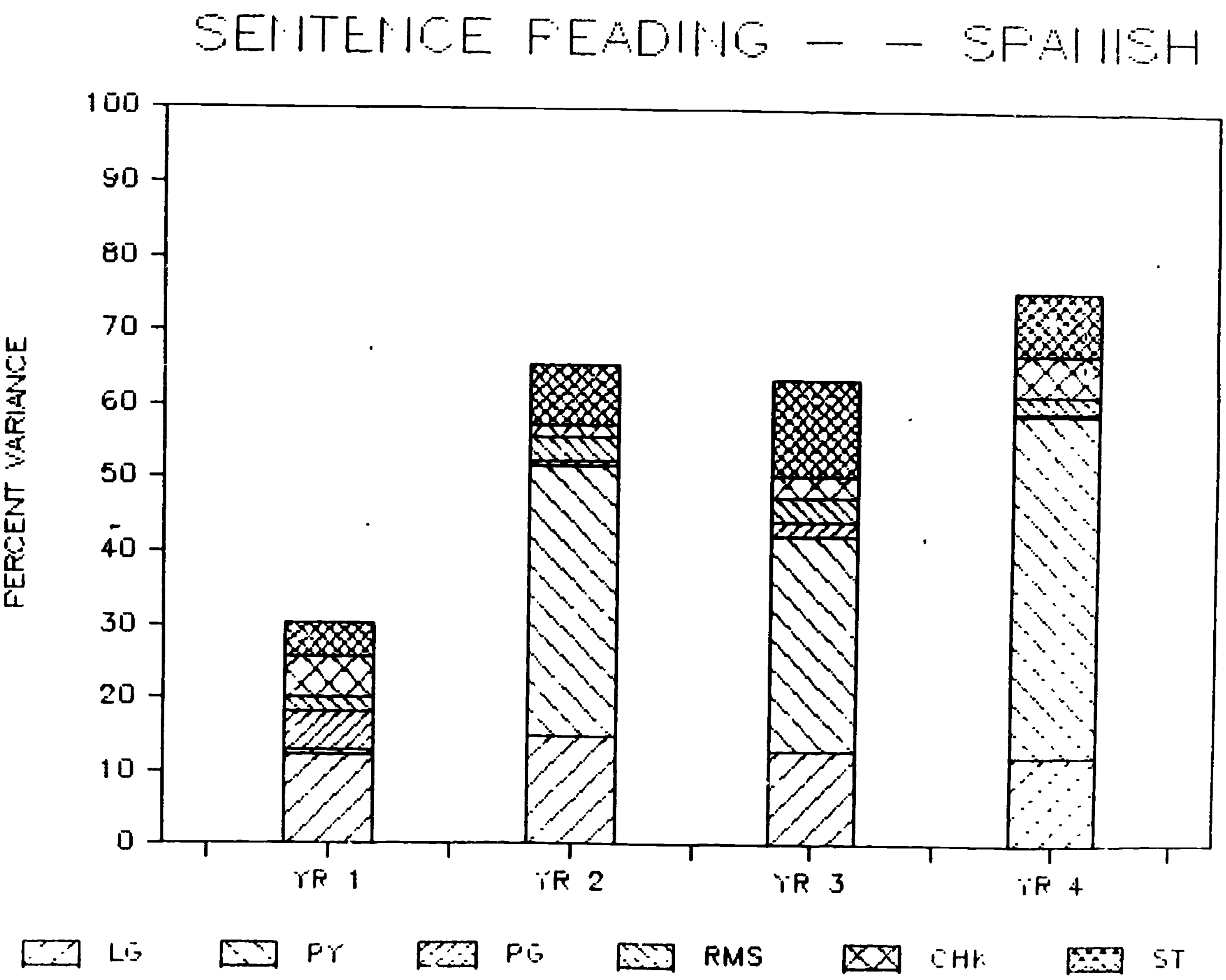


Figure 19. Distribution of explained variance in IRAS-S Sen .e Reading.

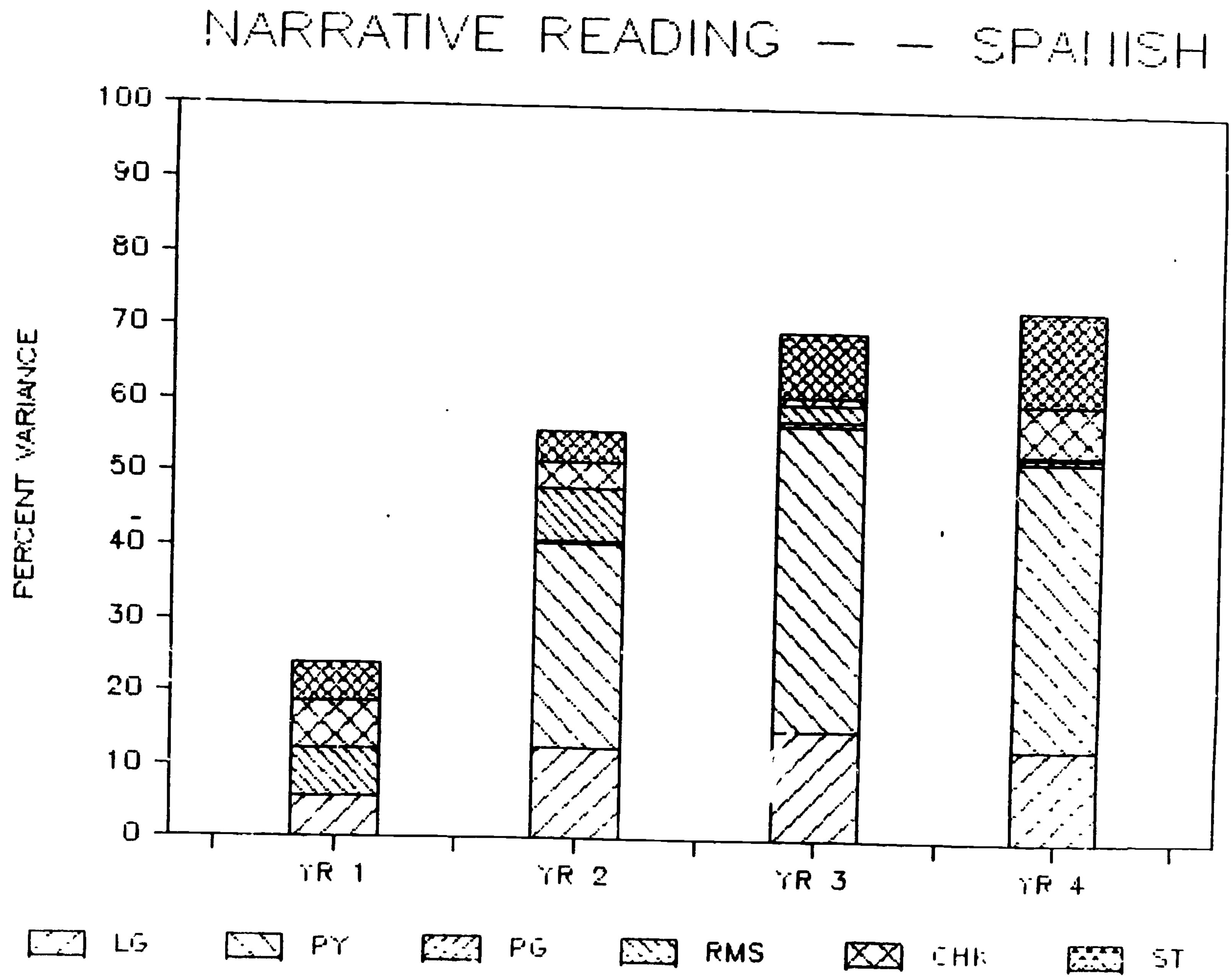


Figure 20. Distribution of explained variance in IRAS-S Narrative Reading Comprehension.

158

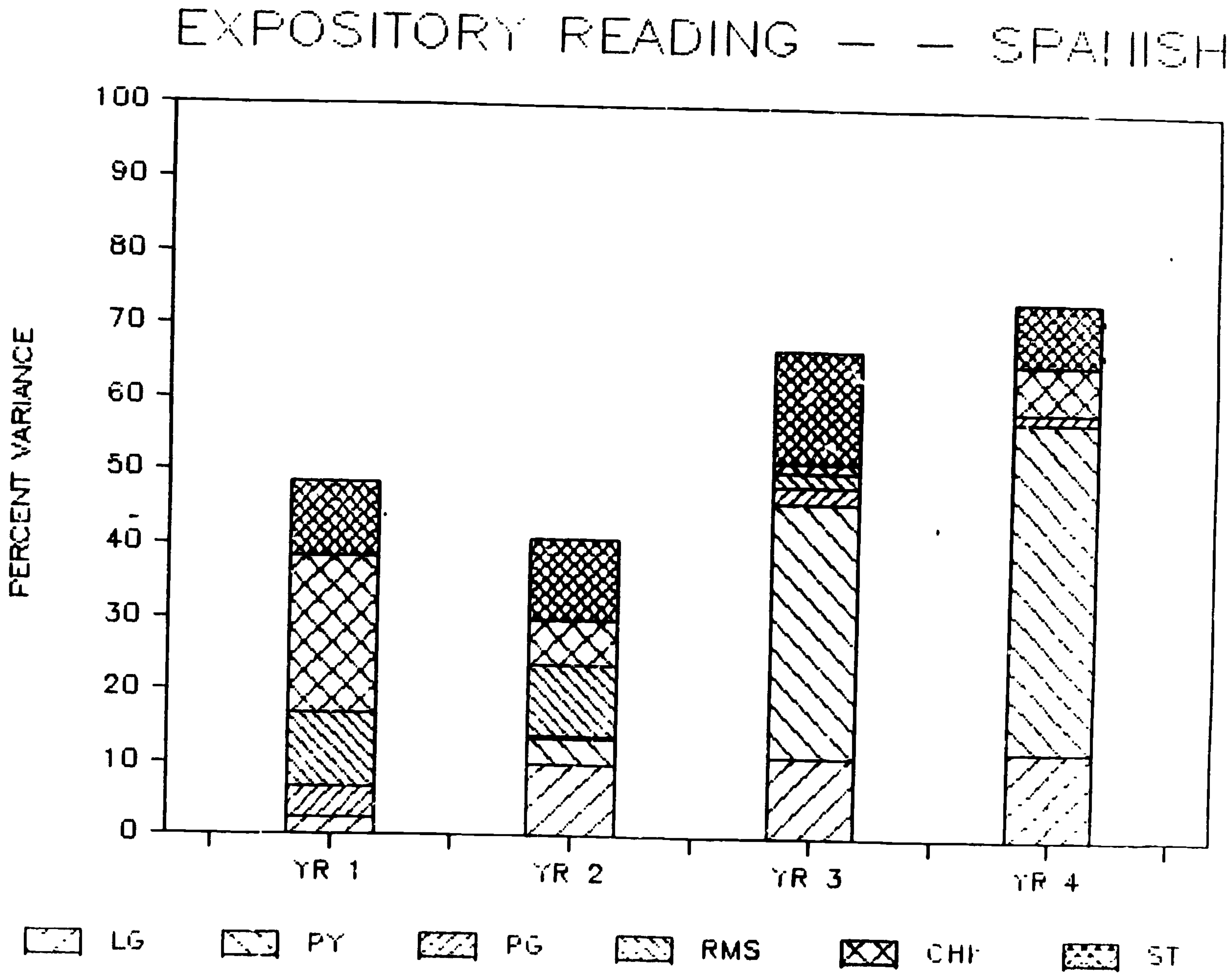


Figure 21. Distribution of explained variance in IRAS-S Expository Reading Comprehension.

predictors do contribute substantially more than would be expected by chance, however, and when their contribution is conditionalized on the variance remaining to be predicted, the effect is quite substantial.

Cluster Pattern

The changing nature of the sample over years makes it somewhat hazardous to average effects over years. With this caveat in mind, some generalizations can be made about the relative contributions of the different clusters of predictors.

For English scales, the average contribution of entry language classification over years ranged from 15 to 20 percent. Previous year's performance added another 17 to 34 percent to the language precursor. Program effects were slight, ranging from 1 to 6 percent depending on the scale. The instructional factors, RAMOS and Checklist, yielded from 10 to 20 percent of additional variance. Attendance effects were small, generally no more than 1 to 2 percent. Finally, the site-specific contrasts gave an average of 5 to 8 percent additional variance.

For Spanish scales, the pattern of effects is comparable to English for most clusters. The exception is the instructional cluster, where the range of contributions is from 4 to 15 percent, smaller than English, and probably reflecting the absence of information about Spanish instruction for many students. Site effects are slightly larger for Spanish scales, ranging from 7 to 11 percent.

For both English and Spanish, the site effects tend to be somewhat larger in Years 3 and 4, even though a single contrast is present in these cohorts and the range of sites is reduced from the first two years of the study. The meaning of this pattern is not clear, but it is persistent.

Specific Findings: Oral Language Scales

The general findings form a framework for considering the more detailed patterns in the regression analyses. These latter will be organized into three categories -- oral language scales, decoding of single words, and reading comprehension -- because of the relative similarities within these three groups.

The oral language scales in IRAS include Vocabulary Definition, and Narrative and Expository Listening Comprehension. The regression analyses for the English versions of these scales are in Tables 46 to 48 and Figures 4 through 6. Language category contributes substantially to each of these scales in all four years, from 10 to 20 percent of the variance in deviations. Alphabet Knowledge in kindergarten does not contribute substantially beyond language; for Vocabulary Definitions, the increment is about 5 percent and is statistically significant, but the increment for Listening Comprehension is negligible.

1035

Previous Year's Achievement matters more in Year 2 and following, when the corresponding IRAS scale from the previous Spring serves as the correlate. The increment in predicted variance from this factor is about 15 percent in Year 2, and increases to double that value in Years 3 and 4, except for Vocabulary Definitions, where ceiling effects may attenuate the relation.

Program (Spanish reading instruction) makes small and inconsistent contributions to predicting the oral language deviates, over and above the precursors. The analysis of partial correlations showed all of the Program effects to be negative in sign.

Instructional factors have a consistent influence on performance over and above the precursors and program factors. Up to this point in the regression analysis, the findings could have been predicted from the correlation analysis. The independent contribution of instruction to deviations is a substantive finding, given the presence of collinearities in the data. Because RAMOS and Checklist factors are partly correlated, it makes most sense to us to consider these variables in combination. The combined contribution of the two clusters is 10 percent in Year 1 for Vocabulary Definitions, 9 percent for Narrative Listening Comprehension, and 7.5 for Expository Listening Comprehension. The corresponding values for Year 2 -- 18, 10, and 13 percent -- are larger than in Year 1, noticeably so in two cases. No effort has been made in these regressions to select instructional factors that are most closely related to performance; the indices were introduced into the regression equation in an omnibus fashion appropriate to the preliminary nature of the analysis. As a result, some factors are being "carried along" that do not carry much weight, which explains the relatively small value of the $\underline{F}$-ratios. It should be remembered that the numerator degrees of freedom for RAMOS and Checklist are 7 and 5, respectively. Significance levels for $\underline{F}$-ratios in this range are $\underline{F}(5, 200, .01) = 3.2$ and $\underline{F}(7, 200, .01) = 2.8$, for purposes of comparison.

Student differences in oral language performance are affected by the instructional variables, over and above the precursor factors. The contribution of these factors in Years 1 and 2, the years in which the entire sample is represented, is substantial as a proportion of total variance, as a proportion of variance remaining after precursors, and as a statistically significant contribution when assessed against residual variance. The pattern continues in Years 3 and 4, with a combined increment of 10 to 30 percent except for Narrative Listening Comprehension in Year 3. Both RAMOS and Checklist factors contribute about equally to explained variance in most of the analyses. RAMOS was introduced first, which gives some advantage to this cluster. In Year 3, the effect of RAMOS on Listening Comprehension was negligible; the zero-order correlations are washed out by Previous Year's Achievement, suggesting that the instructional effects are confounded with entry level in these cohorts.

Attendance has little consistent effect on oral language deviations as an independent predictor. The $\underline{F}$-ratios for Year 2 draw attention, but they miss significance at the .01 level ($F[1, 200, .01] = 6.8$), and account for relatively little variance.

The site contrasts also have little effect on oral language deviations in Year 1; the contribution is only a few percentage points and is not statistically significant. The independent influence of site differences in Year 2 was more noticeable, ranging from 5 to 8 percentage points, and with F-ratios exceeding the .01 criterion for all three measures. Examination of the details of the site effect showed little consistent relation for the three oral language measures; the late order of entry of this factor made interpretation of the beta weights problematic. Site effects in Years 3 and 4 were substantial, even though only a single contrast applied to these cohorts; as noted above, the meaning of these effects is not altogether clear.

The regression analyses for the Spanish versions of the IRAS oral language scales are in Tables 55 through 57 and Figures 13 through 15. Entry language classification contributes substantially to the oral language scales in Years 1 and 2, with values ranging from 20 to 30 percent. Alphabet Knowledge adds virtually nothing to this base in Year 1; previous achievement on the corresponding IRAS scale in Year 2 doubles the percentage of variance accounted for in each of the three scales. The influence of entry language tends to drop in Years 3 and 4 (remember that these are different cohorts), while Previous Year's Achievement has a somewhat greater impact on variability in performance, ranging from 30 to 40 percent of the total variance.

Program effects are noticeable in Year 1, accounting for 6 to 7 percent of the variance in each measure. This cluster of predictors has little influence on performance in later years.

Instructional factors contribute less to the regression equation for the Spanish measures than was true for English. The percentage values for RAMOS and Checklist combined range from 2 to 11 points, without any recognizable pattern over years and scales. As noted earlier, the potential impact of these variables is undercut by the relative infrequency of Spanish instruction in the sites included in the study.

Attendance does not appear to influence performance on any of the oral language scales during any year. The incremental percentages are only 1 or 2 points in most instances, and none of the F-ratios approach significance.

Site contrasts contribute from 5 to 20 percentage points to the variance accounted for. The largest values are in Year 4, where other factors have less impact. This pattern makes sense if one assumes that the role of the school in promoting skill in spoken Spanish is complete, for practical purposes, after the first or second grade, in which case the influence of the community becomes paramount. Interestingly, the contrast between the two border sites, which is the contrast that continues through all four years, is often the largest of the three contrasts in Years 1 and 2.

Specific Findings: Decoding Single Words

Three IRAS scales required the student to decode words in isolation -- Vocabulary Decoding, Letter-sound Decoding (synthetic words), and Letter-sound Spelling. In addition, Sentence Reading comprised a similar task for many students, both because of limited decoding skills, and because the task did not entail any comprehension requirements. Accordingly, we will consider all four scales in this section; the patterns are generally similar.

The regression analyses for the decoding scales are shown in Tables 49 through 52 and Figures 7 through 10. Language Category, the first variable introduced into the regression equation, accounts for 10 to 20 percent of the variance in each scale and during every year, with no clear trends. The pattern is similar to that for the oral language scales in English.

Previous Year's Achievement in Year 1, assessed by Alphabet Knowledge, makes a substantial contribution to explained variance over and above entry language classification, generally by a margin of 10 to 15 percent. This pattern contrasts with the negligible effect of Alphabet Knowledge on oral language scales. This finding gives more precise meaning to the frequently reported correlation between knowledge of the ABC's and acquisition of reading; the effect is linked to decoding and is relatively independent of oral language performance. Previous Year's Achievement in Years 2 and following is much more strongly correlated with the outcome measures, increasingly so from Year 2 through Year 4. By the fourth instructional year, decoding performance at the end of the previous year accounts for 40 to 60 percent of the variability at the end of the following year. The implication is straightforward -- as the student proceeds through the grades, the student's relative standing in decoding becomes fixed in place.

The Program contrasts have little incremental influence on decoding performance. With one exception, the contribution of these predictors is only 1 to 3 percent. For Sentence Reading, the increment is 3 percent in Year 1 and 10 percent in Year 2. Interestingly, the effects of Program contrasts on Reading Comprehension are similar to those observed for Sentence Reading, suggesting that the latter may be influenced by assignment (or nonassignment) to Spanish reading instruction in ways that resemble comprehension effects.

Instructional effects on decoding performance are quite noticeable. The largest effect is during Year 1, which is the grade of greatest importance to decoding instruction in most basal programs. The combined contribution of RAMOS and Checklist ranges from 15 to 1[illegible] percent. RAMOS is the larger contributor; it does have priority of entry, to be sure. This pattern is reversed for Sentence Reading.

The influence of instructional predictors in Years 2 and following is somewhat smaller, though still quite noticeable. The increments range from 5 to 13 percent for the three single-word scales; the effects for Sentence Reading are somewhat smaller, ranging from 2 to 10 percent.

The site contrasts have mixed effects on decoding performance. The contributions range from 2 to 15 percent with no clear pattern; there are both large and small increments for each scale and at each year. The contrast between the two sets of border sites is again often the largest effect.

The regression analyses for Spanish decoding scales are presented in Tables 58 through 61 and Figures 16 through 19. Entry language classification has a relatively consistent effect on decoding, generally ranging between 10 and 15 percent with a couple of exceptions. In contrast with the situation for English, Alphabet Knowledge made no contribution to decoding performance in Year 1 for Spanish. In Year 2 and subsequent, previous performance made a substantial contribution to prediction of deviations at the end of the following year. The increments range from 30 to 60 percentage points, typically increasing from Year 2 through Year 4.

Program contrasts have little influence on the variability in decoding scores. While these contrasts are correlated with the scores, Language Category and Previous Year's Achievement absorb most of this relationship.

The pattern of influence by instructional factors on Spanish decoding resembles that for English decoding. The largest effects are observed during Year 1, where the increments for the three single-word scales range from 13 to 18 percent using the combined scores from RAMOS and Checklist indices. Sentence Reading was not so strongly influenced, reflecting our observations that Spanish instruction stressed the decoding of single words in isolation, with little emphasis on fluent reading of connected text or on comprehension. Performance in Years 2 and following was not incrementally affected by instructional variables to any noticeable extent. The percentage of incremental change in total variance accounted for ranges from 3 to 12 points, generally toward the low end of the range and seldom attaining statistical significance. Again, this pattern makes sense given the tendency to emphasize phonics in first grade, and the fact that relatively little instruction in Spanish occurred after first grade.

Attendance had more influence on Spanish decoding scores than most other IRAS scales, especially during the second instructional year when the variable was statistically significant on two of the four scales. At best, this factor still remains only a modest influence on performance.

Site contrasts accounted for about 5 to 10 percent of the residual variance in the outcome measures over scales and years.

1039

Again it is the comparison between sets of border sites that appears as the most consistent source of variance.

Specific Findings: Reading Comprehension

The regression analyses for the English Narrative and Expository Reading Comprehension scales are displayed in Tables 53 and 54, and in Figures 11 and 12. Entry language classification yields an increment of about 10 to 15 percent as a baseline except in Year 3, where the effect is about twice as large. Alphabet Knowledge adds another 15 percent to this base in Year 1, and Previous Year's Achievement correlates contribute increasing increments from Years 2 through 4, ranging from 20 to 45 percent, respectively.

Program effects account for 5 to 10 percent of the systematic variance from Years 2 through 4; the effect in Year 1 is negligible. This pattern differs from that observed for other measures, and is not immediately explainable. One would expect the larger effects of Program to come during the first year of instruction.

Instructional factors contribute a substantial amount of predicted variance to both scales at all years, an increment ranging from 10 to 15 percent with only two exceptions. There are no obvious trends; both RAMOS and Checklist factors share in the systematic variance about equally.

Site effects on Reading Comprehension are negligible in Year 1. They increase from Year 2 through Year 4, ranging around 5 to 15 percent.

Regression analyses for Spanish Reading Comprehension are shown in Tables 62 and 63, and in Figures 20 and 21. The patterns for these measures are quite unlike most of the others, in large measure because of the relatively low levels of performance on these scales, especially in the first instructional year or two. Neither Language Category nor Previous Year's Achievement influenced variability in Year 1 to any noticeable extent. From Year 2 through 4, Language Category contributed 10 to 15 percent to the systematic variance. Previous Year's Achievement was a more substantial determinant of performance during this time, with increments ranging from 30 to 40 percent with the exception of Expository Reading Comprehension in Year 2.

Program contrasts did not have any independent effect on performance, as was true for decoding measures. Instruction did make a difference, however, during Years 1 and 2. The indepdendent contribution of RAMOS and Checklist factors to variability in comprehension measures ranged from 10 to 15 percent during these years, an effect both substantial and statistically significant. The contribution of instructional indices in the last two years was negligible, as might be expected from the relative absence of comprehension instruction during this time. Comprehension was not a primary focus during the first two years, but students gained in decoding skills, which led to some advances in reading comprehension of relatively simple passages.

Attendance was not a major contributor to comprehension performance, reaching statistical significance in only one of the eight comparisons. Site contrasts contributed from 5 to 15 percent additional variance to the regression equation, with no clearcut trends. Again, the differences between the border sites appeared to be generally more important than the other contrasts.

DISCUSSION

In review, the analytic approach applied in integrating this study's complex data sources called for examination of (a) the yearly correlations between individual IRAS subscale deviates and the set of summary variables reflecting student entry characteristics and subsequent instruction and (b) a series of regression analyses designed to determine the amount of variability contained in the sets of IRAS deviates that could be explained by the student entry and subsequent instruction variables. Generally, the findings suggest that 75% to 95% of the IRAS deviate variability was associated with the predictor set, and that the strongest predictive relations were for kindergarten entry language skill, performance during the previous year, and instruction. In discussing these findings, we will treat precursor skills first.

Precursor Skills

As noted throughout this report, entry oral language and prereading skills were associated with reading achievement. When compared with children with less well developed skills at entry, children with relatively better developed oral language and prereading skills at entry were better able to take advantage of the instruction offered and to maintain their relatively superior level of attainment in reading throughout the primary grades.

Entry English language skills have pervasive and lasting effects on English reading achievement throughout the early elementary grades. Recall from Volume 5 that while the oral English growth rate of the overall sample proceeded at a rate above the expectations of the growth track model, that of the Low English entry students showed a greater rate of growth than that of the High English entry group but did not converge with that of the High English entry group until late fourth grade. This suggests that the acquisition of "school-related" skills in a second language takes time. An important question for educators is, "What are effective intervention strategies for ensuring academic progress during the years while these children are in the process of gaining the necessary proficiency in English?"

An additional important challenge for educators is finding means to ensure, and perhaps accelerate, language and reading growth of students who at entry into school are deemed by their teachers to have relatively low level verbal skills. These are the children who got off to a slow start in school, gained somewhat less than a year of growth for a year of instruction, and fell further behind their more academically-prepared peers as they progressed through the early elementary grades.

Nominal Instructional Program

The study examined the degree to which the number of years students were enrolled in a Spanish reading program could account for reading achievement within each of the instructional years. For Spanish literacy, enrollment in Spanish reading programs was generally positively related to reading achievement during the early grades, but this relationship became negligible in the later grades. Children who are placed in these programs are generally those who are deemed by the schools to be limited in their English skills and to have stronger skills in Spanish than in English at the point formal reading instruction is begun, usually in first grade. These children remain in Spanish reading programs until they (a) reach a predetermined level of oral proficiency in English and (b) have attained a specified level of reading in Spanish and/or perform at or above a specified percentile score on a standardized test of reading achievement in English (usually the 40th percentile). In this study, some students in these programs received Spanish reading instruction for one year before being transferred to English reading; others remained in Spanish reading for two, three, or four years, with most being transferred to English reading by the end of third grade. Once transfer to English occurred, no further reading instruction in Spanish was provided, except during a brief "transition" period in some schools. With such criteria for transfer, the few students who remained in Spanish reading programs beyond the third grade were likely to have been children who were having trouble learning to read, since the oral English skills of most of the students by third grade exit tended to meet or exceed the oral English criterion for transfer. Thus, the failure to find a relationship between Spanish reading assignment and Spanish reading achievement in the later grade levels is not surprising as the highest achievers have most likely been transferred out of the Spanish reading program.

While acquired English literacy skills were found to be generally negatively associated with numbers of years of enrollment in Spanish reading programs, there was some indication of relatively superior English literacy skills at fourth grade exit for those students with longer (longitudinal) enrollments in such Spanish reading programs. Although the sample was limited for this instructional year, this trend in the data raises some interesting questions. For children who begin initial reading instruction in Spanish, is there a threshold level that must be reached in Spanish reading for the benefits of such instruction to positively affect growth in English literacy? If so, does it correspond to the level of literacy that monolingual children normally achieve by the end of third grade? Are children in transitional bilingual education programs, where criteria for transfer to English reading is strongly tied to English performance, being kept in Spanish reading programs sufficiently long for them to attain the requisite literacy skills in Spanish? Does the time frame of this study, kindergarten through grade four, capture the long-term effects of initial reading instruction in the non-English home language? These and related questions merit attention as they are central to the current controversy surrounding transitional bilingual education.

Quantity and Quality of Instruction

Of the many factors that impact on student progress in reading, instruction is the one factor for which the schools have primary responsibility and over which they have the most control. Therefore, identifying instructional patterns that are associated with success and failure, both in the early stages of reading instruction and in subsequent years, is a critical issue surrounding improvement of practices for all children. In this regard, the finding that instructional variables make substantial contributions to achievement (in each of the three domains of oral language, decoding, and reading, and in both English and Spanish) beyond precursor effects is important -- instruction does make a difference.

The classrooms in this study exhibited several of the characteristics of effective instruction, and for the students (in the aggregate) such instruction produced approximately a year of growth for a year of instruction in English reading comprehension as measured by performance based tests.

While similar in many ways, variation was noted among the classrooms on the quality of the dimensions of instruction assessed in the study. This suggests that to ensure effective instruction of all students, certain instructional dimensions need to be strengthened. Staff development should aim toward training teachers to (a) monitor their own use of language in the classroom and to provide instructional activities that make strong formal language demands on students; (b) make optimal use of textual materials, favoring these over non-textual materials in both direct instruction and independent work; (c) increase instruction in word meaning and the higher-order comprehension skills and to strengthen such skills through making this instruction explicit; and (d) evaluate the decoding needs of their students and to tailor their instruction on decoding to the identified needs, making such instruction explicit and limited to appropriate amounts. In addition, the practice of grouping students for instruction needs careful consideration, not only in terms of optimal size but also in terms of student membership, permanency of the group once formed, and instructional treatment provided.

1943

Site Characteristics and Reading Achievement

The sites included in the study were selected to achieve variation on several dimensions (e.g., size, socioeconomic status, degree of urbanicity, concentration of Hispanic students, characteristics of the reading program). Given differing contextual environments, site differences in language and literacy development could be expected. For English, site contrasts in such development were relatively isolated, suggesting that schools were adjusting schooling practices and instruction to accommodate the needs of the local school population. Spanish literacy, on the other hand, was more advanced at certain of the border sites where substantial non-school support for Spanish was available.

Factors outside of the school play an important role in maintaining and/or fostering development of the non-English home language. Prominent among these are locale and the extent to which the language is used in the community and the wider environment, as well as the role of the home language in the affairs of the home and of the community; attitude of the student and others toward the maintenance of Spanish; and the extent to which written materials and formal usage are available to the students in the home language.

Without strong support from the home and the community, students in transitional bilingual education programs are not likely to achieve high levels of literacy in Spanish. Indications are, however, that these programs can, and are, promoting English literacy for all students. In this study, the students on the average were acquiring English oral language skills at the rate expected and were gaining in English literacy at or near a year of gain for a year of instruction, depending upon the type of reading measure used. Further, slightly more than half of the students were reading in English at grade level expectations by the end of second grade. Are these realistic expectations for schools to hold for students from non-English language backgrounds who at entry into school are limited in their English skills? Are growth rates of these youngsters similar to those generally reflected by mainstream school children under current schooling practices in public school systems in the United States? Do the growth rates of these youngsters compare favorably with those of their monolingual peers in the same school?

References

Barr, R., & Dreeben, R. with Wiratchai, N. (1983). How schools work. Chicago, IL: University of Chicago Press.

Calfee, R. C., Mace-Matluck, B. J., & Hoover, W. A. (1982, December). Teaching reading to bilingual chilldren: Effects of interaction of cognitive style and reading methodology on the reading achievement of bilingual children (Interim Report, Contract No. 400-83-0007). Washington, DC: National Institute of Education.

Calfee, R. C., & Piontkowski, D. (1984). Design and analysis of experiments. In P. D. Pearson (Ed.), Handbook of reading research. New York, NY: Longman.

Cronbach, L. J., & Snow, R. E. (1977). Aptitudes and instructional methods. New York, NY: Irvington Publishers.

Hoover, W. A. (1983). Language and literacy learning in bilingual instructin: Cantonese site analytic study (Final Report, Contract No. 400-80-0043). Washington, DC: National Institute of Education.

Rogosa, D. R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. Psychological Bulletin, 92, 726-748.

1045

Final Report

TEACHING READING TO BILINGUAL CHILDREN STUDY

Volume 8

Executive Summary

Southwest Educational Development Laboratory

This report is one of a series produced for the

TEACHING READING TO BILINGUAL CHILDREN STUDY

by the Southwest Educational Development Laboratory:

This document was produced under Contract No. 400-83 0007 from the National Institute of Education, U.S. Department of Education, to the Southwest Educational Development Laboratory. However, it does not necessarily reflect the official views of either the Institute or the Department. Information about the study can be obtained from the National Institute of Education and from the following persons from the Southwest Educational Development Laboratory who have served as Co-Principal Investigators for the Study: Betty J. Mace-Matluck and Wesley A. Hoover.

Final Report

TEACHING READING TO BILINGUAL CHILDREN STUDY

Volume 8
Executive Summary

Betty J. Mace-Matluck, Wesley A. Hoover
Robert C. Calfee

Document BRS-84-R. 1-VIII

Preston C. Kronkosky, Executive Director
Southwest Educational Development Laboratory
Austin, Texas

November 1984

There were many individuals and institutions who contributed to this research effort. We wish to express our sincere gratitude to the parents, students, and school personnel who provided the necessary data from which this study is derived.

In addition, several other individuals made valuable contributions to the study, for which we are indebted: Robert C. Calfee, Sylvia C. Peña, and Blanca de Alvarez.

And finally, we wish to thank the local data collectors at the school sites, many of whom remained with the study throughout its duration: Ramiro Barrera, Beatrice Cantú, Irene Cavazos, Carolyn Cruz, María de Obregón, Gloria de Torres, Gigi Galván, Olga Hernández, Irene Mendez, Guadalupe Treviño, Rosalinda Villarpando, and Gloria Villarreal. Their patience, dedication, and hard work helped make this study a reality.

Betty J. Mace-Matluck
Wesley A. Hoover

TABLE OF CONTENTS

1251

PREFACE

In June 1978 the National Institute of Education (NIE) funded the Southwest Educational Development Laboratory (SEDL) to conduct a longitudinal study on the Teaching of Reading to Bilingual Children. Educators and policymakers alike have long recognized that the ability to read is essential for success in school, in work, and in life; yet many children from second-language backgrounds have trouble learning to read in schools today. The majority of these youngsters are from Spanish-language backgrounds and from low income families. Special programs designed to meet the needs of these children are provided in schools, but there is limited research evidence to guide the development, evaluation, and implementation of these programs. This study is intended to provide information that will result in greater insights into what constitutes a favorable learning environment for children from Spanish-language backgrounds, what instructional sequences and events promote successful and efficient learning of literacy skills, and what the language and literacy outcomes of current schooling practices are for a large sample of these youngsters.

The study was conducted during the years of 1978 through 1984. It is a comprehensive longitudinal investigation of the development of reading skills from kindergarten through fourth grade for a representative sample of more than 350 children from bilingual backgrounds, and for smaller samples of children who, on entry into school, were monolingual in English or Spanish. In this "natural variation" study, teaching and learning were carefully documented in field settings at the several sites.

The goals of the study were to (a) describe variations in both English and Spanish language ability of students living in bilingual communities, (b) document prevailing practices in reading instruction for bilingual students, and c) investigate the relations between the instructional program and student achievement for students with differing entry profiles.

Description of the Study

Surveys of the general and school populations reveal an increase in the number of students whose language resources are not an ideal match to the language of the school. An important question for educational practice and policy centers around the school's responsibilities in this situation. Bilingual programs, English-as-a-Second-Language classes, classroom aides, and "sink-or-swim" approaches can all be found in practice today. From limited evidence now available, none of these techniques has emerged as the one best system.

Hispanics make up the largest and fastest growing school-age population today. The demographics for some states show that over the next decade they may constitute as much as a third to a half of the population. In the state of Texas at present approximately one third of the school children are from Hispanic backgrounds (approaching one

million). They are found in virtually ever school district in the state. Many of the school districts in the southern portion of the state serve school populations of which 75% to 99% of the children are from Spanish-speaking backgrounds and, on entry into school, are often limited in their ability to speak English and to profit from instruction in that language This population is not restricted to the border areas, however. Large centers in the state report as much as 20% of their school populat[illegible] rom Hispanic backgrounds, with a concentration of some 80% to 90% [illegible] certain of their schools.

It is well documented that, in general, children from Spanish-speaking backgrounds, for whatever reason, often encounter difficulty in our nation's schools; they do more poorly on standardized tests than does the general school population, and their dropout rate is high. Bilingual education, in which students are given instruction partially through the home language until they have attained sufficient proficiency in English to benefit from English-medium instruction, has been the principal approach recommended by the Office for Civil Rights to ensure access to equal educational opportunity for these children. Although many individual programs have had considerable success in improving the academic performance of language-minority students, it has not been demonstrated that these programs generally are reducing inequality of educational opportunity on the large scale that was envisioned.

Growth in reading comes about for most youngsters through formal classroom instruction. Understanding the development of reading, and knowledge of the critical variables that determine success or failure, depends on a careful examination of the instructional program -- not just the label over the classroom door, but the program as actually implemented by the classroom teacher.

Educators have raised several issues about the most effective way to help bilingual children become proficient readers of English. These include (a) valid assessment of the student's ability in the languages of the home and of the school, (b) the optimal balance of formal instruction in both languages, (c) the most effective transfer from one language to the other, and (d) bilingual support within the classroom environment. A major thesis of the Teaching Reading to Bilingual Children study is that addressing these issues (and others) requires a comprehensive and ecologically-valid investigation of the linkage between the child's language and the language of instruction.

Design of the Study

To achieve the objectives of the study, considerable attention was given to the selection of schools, teachers and students, to the instruments for assessing language and reading achievement, and to the methods for evaluating the classroom instruction. Each of these topics is discussed briefly below.

Schools, Classes and Teachers

Twenty schools and 200 teachers from six school districts participated in the study. Included are variations in the nature of the reading program (a range from phonics-oriented to meaning-based), classroom organization (some self-contained, others team-taught), and grade structure (the range of grades in the individual school and the extent of cross-grading both vary). The schools differed in size, SES, urbanicity, locale, and makeup of the student body (from medium to high concentration of bilingual students).

Student Cohorts

The study was undertaken in four cohorts or "waves" of students. Three of the cohorts consisted entirely, or in large part, of bilingual students. The first cohort was small (N=40) and of limited generality; the second was somewhat larger (N=80) and covered a slightly broader array of contexts. The third cohort which was both larger (N=200) and broader in its generality, incorporated a number of procedural improvements based on previous experience in the study and included a monolingual English-speaking sample. The fourth cohort consisted of a relatively small sample (N=60) of monolingual Spanish-speaking students.

All of the bilingual sites were from the state of Texas, as were the monolingual English-speaking students. The monolingual Spanish-speaking students were from one site in Northern Mexico.

The original design of the study called for each student to be assessed and observed from entry to kindergarten through exit from third grade. By covering the full range of the primary years, we would be able to examine the transition from "learning to read" through "reading to learn." For students in programs where the initial stages of reading were in Spanish, we also considered it important to determine the transition to competence in English reading.

The original design was in fact implemented for the first two cohorts; some of the students were tracked from first through fourth grade, but most followed the intended design. Due to limited funding in the later stages of the study the last two cohorts could not be followed for the full four years that were originally intended. The bilingual and monolingual English samples from the Texas sites were observed from kindergarten through second grade, and the monolingual Spanish samples from the site in Northern Mexico were observed from first through third grade (the program did not provide a kindergarten).

The monolingual samples were incorporated in the design to aid in validating the instruments for student assessment. Both the English and Spanish cohorts are small and not selected to be fully representative of monolingual populations. Data from these samples will be presented in Volume 3, as part of the discussion on the adequacy of the instruments for measuring growth. The study was designed to study the course of reading in bilingual students, not as a basis for comparing these students with monolingual youngsters. Accordingly, comparisons

between the various samples will not be made in this report, nor do we recommend that others attempt such comparisons.

Language Assessment

Several types of data were collected for each student on English and Spanish proficiency. Each year, early in the Fall and again in the Winter and Spring, teachers rated their students' language skills. Oral language proficiency tests were administered in the Fall of each year. Finally, audiotaped speech samples were obtained monthly on a rotating schedule in three settings: in the classroom, on the playground, and in the home.

Reading Assessment

Several instruments were used to measure reading achievement. Standardized test scores (mostly English) were collected yearly. More detailed information was obtained from a battery of individually-administered "performance based tests" in both English and Spanish. In kindergarten, the Stanford Foundation Skills Test was employed to measure the child's pre-reading skills. From the end of first grade on, the Interactive Reading Assessment System was administered during the Spring of each school year. This instrument provides independent measures of the student's skills in decoding, word meaning, fluency in oral reading, and comprehension. Finally, informal reading inventories were administered throughout the school year.

Classroom Observations and Teacher Interviews

Project staff conducted monthly observations of the reading instruction in each classroom and interviewed the teachers quarterly about their instructional plans. The observation instrument documented staffing patterns, grouping and organization, time allocation, the language of instruction, the character of instruction, the materials and procedures used, and the response of the students. The interviews focused on the teacher's general instructional objectives, as well as the objectives for individual target students. Taken together, these two instruments yield a rich characterization of the classroom environment for the target students.

Student Entry Variables, Classroom Factors, and Reading Achievement

The primary goals of the analyses were to identify the general relationships that characterize variation in these factors and to look for underlying regularities that are associated with success and failure, both in the early stage of reading instruction and in the year-to-year variations.

Documents

This report is one of a series of eight documents contained in the Final Report submitted to the National Institute of Education. A com-

plete list of these documents is provided on the inside of the cover of this report.

The study was a collaborative effort among a number of individuals and institutions. All members of the research team contributed to the thinking, planning, and writing of this series of documents, however, the individual whose name appears first in the list of authors was responsible for preparing the particular document.

Betty J. Mace-Matluck
Wesley A. Hoover
Co-Principal Investigators

Austin, Texas
November 30, 1984

INTRODUCTION

Under contract with the National Institute of Education, the Southwest Educational Development Laboratory (SEDL) conducted a comprehensive, six-year longitudinal investigation of the development of language and reading skills during the primary grades for a representative sample of more than 250 Texas children from bilingual backgrounds, and for smaller samples of children who are monolingual in English or Spanish. The research was initiated on June 1, 1978, in response to needs expressed at the national and regional levels for information that could assist policy makers and practitioners in planning and delivering effective language and reading instruction to children from complex language backgrounds.

Designed to examine the relations between current schooling practices and the language and reading achievement of a large sample of low income Hispanic children who began their initial schooling in bilingual classrooms, the study sought to provide information that could result in greater insights into (a) what constitutes a favorable learning environment for children from Spanish-language backgrounds, (b) what instructional sequences and events promote successful and efficient learning of literacy skills, and (c) what the language and literacy outcomes of current schooling practices are for these youngsters.

Surveys of the general and school population reveal that Hispanics make up the largest and one of the fastest growing school-age populations today. The demographics for some states show that over the next decade they may constitute as much as a third to a half of the population (Hispanic Policy Development Project, 1984; O'Malley, 1982). In the State of Texas, approximately one-third of the school children are from Hispanic backgrounds; 50% of the current kindergarten population in the state is Hispanic. Hispanic children are enrolled in virtually every school district in Texas, with many of the school districts in the southern portion of the state serving school populations of which 75% to 99% of the children are from Spanish-speaking backgrounds and on entry into school are often limited in their ability to speak English and to profit from instruction in that language. In addition, certain of the large urban centers in the state report as much as 20% of their school population from Hispanic backgrounds and a concentration of some 80% to 90% Hispanic in some of their schools.

It is well documented that, in general, children from Spanish language backgrounds, for whatever reason, often encounter difficulty in our nation's schools -- they do more poorly on standardized achievement tests than does the general school population, and their dropout rate is higher. However, this population is not a homogeneous group. Differences are found in their degree of bilingualism, immigrant versus long-term residence, country or region of origin, socioeconomic conditions, mobility, the way in which English and Spanish are used in the various domains of life and thought, and in their experiences in dealing with print both in and out of school. Of those who experience difficulty in learning to read and write in school, an overwhelming majority are from low-income families and from environments where Spanish is widely used both in the home and in the community. The Teaching Reading to Bilingual Children Study, therefore, focused on Spanish-speaking children from low-income families in Texas.

Growth in reading comes about for most youngsters through formal classroom instruction. Understanding the development of reading and knowledge of the critical variables that determine success or failure depend on a careful examination of the instructional program -- not just the label over the classroom door, but the program as actually implemented by the classroom teacher.

Educators have raised several issues about the most effective way to help bilingual children become proficient readers of English. These include (a) valid assessment of the student's ability in the languages of the home and of the school, (b) the optimal balance of formal instruction in both languages, (c) the most effective transfer of skills from one language to the other, and (d) bilingual support within the classroom environment. A major thesis of the Teaching Reading to Bilingual Children Study is that addressing these and other issues requires a comprehensive and ecologically-valid investigation of the linkage between the child's language and the language of instruction.

The research, a "natural variation" study in which teaching and learning were carefully documented in field settings at six sites, was guided by the following goals:

1. To describe variation in both English and Spanish language abilities of students living in bilingual communities.

2. To document prevailing practices in reading instruction for bilingual students.

3. To investigate the relations between the instructional program and student achievement for students with differing entry profiles.

To accomplish the goals of the study, more than 300 children in 20 schools in 6 school districts and taught by more than 200 teachers were tracked from kindergarten through second or third grade (fourth grade in some cases) -- a critical period for the development of literacy.

The study examined the children's language on entry into school and thereafter. Standardized test data were collected and examined, as were other more detailed sources of language and reading data. Systematic observation was carried out in the classroom. Information was gathered about the teachers' instructional plans, and the nature of the instructional program was carefully documented.

The students' entry skills and subsequent achievement in the various components of language and reading were examined, and the students' rates and patterns of growth were investigated in relation to the instruction each student received.

The primary analyses of the data from the study aimed toward four basic areas:

1. Precursors. Descriptive information using validated precursor profiles typically found in bilingual children on entry into schools throughout Texas.

2. Instruction. Class-level descriptions of the approaches used to teach reading to children from bilingual backgrounds in the state.

3. Achievement. Development and validation of a set of longitudinal achievement indices that could be used to assess growth in the various components of reading in English and Spanish.

4. Linkage. Development and validation of a set of procedures for measuring the linkage between reading achievement on the one hand and precursor and instructional indices on the other hand, taking into account the possibility of interactions between precursor profiles and response to type of instruction.

The results of the analyses in these four areas (Precursors, Instruction, Achievement, Linkage) were presented in previous volumes of the final report. Those volumes have provided a general introduction to the study (Volume 1); described the overall design (Volume 2); discussed the methodology used in the analysis of the data (Volume 3); presented a discussion of the instruments and the data obtained from them for the bilingual sample, reporting descriptive and summary statistics for each of the data sources within each of the major elements of the data structure (Volume 4, language; Volume 5, pre-reading and reading; Volume 6, instruction), and explored the linkages between the different sources of information contained within the data base (Volume 7).

This document, Volume 8: Executive Summary, provides an overview of the study, a summary of the findings, and a general discussion and interpretation of the results. The remainder of this document is organized around six major sections: Research Design, Oral Language, Instruction, Reading, and Integration of Data Sources.

RESEARCH DESIGN

To achieve the objectives of the study, considerable attention was given to the selection of schools, teachers, and students; the instruments for assessing language and reading achievement; the methods for assessing classroom instruction; and the data analysis plan. Each of these topics are discussed briefly below.

Sample Description

Having determined the goals of the study and that low-income Hispanic children enrolled in bilingual programs in the State of Texas would be the primary focus of the study, a <u>purposive</u> rather than a probability sampling procedure was selected. The sampling plan included sampling at various units of analysis: region, school district, school, teacher/classroom, and student.

The general approach employed was to start at the highest level of the chain with the selections of regions, and proceed to sampling at lower levels, using data available at each point to establish fixed categories from which samples were to be taken. Data compiled by the Texas Education Agency and previous work carried out at SEDL suggested that two or three general types of bilingual education programs could be identified with two or three reading approaches nested within, or across, the bilingual programs.

Site Selection

The initial selection of sites was based on the division of the state into geographical regions that took into consideration a combination of regional, political, socioeconomic status, language, and degree-of-urbanicity variables. Four geographical regions were then selected:

Central Texas -- a region that is both urban and rural and contains a number of bilingual programs.

Texas Border Area -- rural, low socioeconomic status, substantial numbers of Spanish dominant students.

North Central Texas -- large urban area, largely monolingual English, middle-sized cities, poor and middle class.

Northern Mexico -- monolingual Spanish, rural and small and middle-sized cities, poor and middle class.

The areas in Texas constituted the primary region from which the bilingual sample was drawn. The monolingual samples were drawn from the Northern Mexico region and the Central and North Central Texas regions.

Selection of School Districts

Within each region, four to eight school districts were identified for potential inclusion in the study. Ultimately six districts were selected that were as broadly representative as possible on the variables of interest (size, socioeconomic status of local community, degree of urbanicity, nature of the bilingual program, degree of variability of schools and teachers within the district, and willingness to cooperate in the study).

Selection of Schools

Schools were selected within each of the districts on the basis of such variables as, but not limited to, the nature of the bilingual program, nature of the reading program, and character of the school organization (e.g., multiple grading, team teaching, open classroom configurations, and individually-guided education programs). Fourteen schools were selected from which the student sample was initially drawn. As the students moved through the grades, six additional schools were involved.

Selection of Teachers/Classrooms

Data available at the district and school level were used in the selection of teachers. Variables considered in teacher selection, as students entered the study initially, included number of years of experience, specialized training in reading and bilingual education, number of years at the present school, and qualifications and role functions of the teacher aides.

As students moved on to the next grade, they were often dispersed throughout all appropriate classes that were available in their school at that grade level, as the study had no control of student placement beyond the students' initial year in the study. Initially, the student sample was assigned to 26 homeroom classes. However, because of team-teaching and other organizational approaches, 37 teachers constituted the initial teacher/classroom sample.

Selection of Students

The students' language and reading skills were assessed with a variety of instruments, and their instructional programs and classroom instruction were observed and documented. For some purposes, all of the students in a class were tested with certain instruments; for other purposes, the instructional program of the entire class was observed. In addition, a target subsample of 10 students was selected in each class for a more detailed, longitudinal "case study" examination.

The primary factors for the selection of target students within a classroom included sex, language status, and an index of cognitive style.

In sum, the five Texas sites selected represent a cross-section of school districts typically found in the state and reflect variation in size, socioeconomic status, urbanicity, locale, and makeup of the student body (from medium to high concentration of bilingual students). They also reflect a high degree of diversity in terms of curriculum used, organization for instruction, criteria and practices for transition from Spanish to English reading, and instructional emphasis. Thus, the naturally occurring variations necessary to the design of the study were found in the sites included in the study.

Cohort Plan for Longitudinal Investigation

To achieve the purpose of the study, it was most desirable to track the target students from entry into kindergarten through the end of fourth grade. The growth and development that are the focus of the study normally takes place over this time peri..., and a cross-sectional design would have been altogether inappropriate.

For practicality, what was planned and carried out was the selection of cohorts (groups) of relatively modest sample size that were tracked for varying periods of time in successive waves. Four student cohorts of differing sizes entered the study during the course of a three year period. Each cohort of students was tracked from their entry into the study through the last data collection year (1982-1983) or until they exited fourth grade. The selection procedure yielded a subsample of 380 students distributed among the cohorts as follows:

Cohort 1: Year 1. Site 0 - 2 schools, 4 classrooms, 40 target students (20 K; 20 F).

Cohort 2: Year 2. Site 0 - 2 additional classrooms, 20 students (K); Site 1 - 1 school, 2 classrooms, 20 students (10 K; 10 F); Site 2 - 1 school, 4 classrooms, 40 students (20 K; 20 F).

Cohort 3: Year 3. Site 3 - 1 school, 3 team teaching units, 80 bilingual students, 10 monolingual English (all K students); Site 5 - 7 schools, 11 classrooms, 80 bilingual students, 30 monolingual English (all K students).

Cohort 4: Year 3. Site 4 - 2 schools, 4 classrooms, 60 monolingual Spanish students (all first graders).

Since the students entered in successive waves, with most of the students entering during their kindergarten year but with some entering at first grade, certain of the students were tracked for five years (K-4); others for four years (K-3 or 1-4); yet others, who were the majority, were tracked for three years (K-2 or 1-3, the latter being the case of the Northern Mexico sample).

The monolingual samples were incorporated in the design to aid in validating the instruments for student assessment. Both of these samples were small and were not selected to be fully representative of monolingual populations. The study was designed to examine the course of reading development of bilingual students, not as a basis for comparing these students with monolingual youngsters. Accordingly, comparisons between the various samples were not made, nor is it recommended that others attempt such comparisons.

Instrumentation and Data Collection

Data were collected in accordance with a data collection schedule that was prepared each summer and distributed to data collectors prior to the beginning of the school year. Two major sets of instruments were administered each year, one assessing student characteristics and academic performance, and a second providing information on teacher characteristics and classroom instruction. A brief discussion of these data sources is given below; more detailed descriptions of the primary measures are provided in subsequent sections of this report.

Student Characteristics and Academic Performance

For students, the study's primary interest focused on language and literacy growth. The instruments employed in the assessments of these skill areas are described below.

Language Assessment

Several types of data were collected for each student concerning their oral language proficiency in both English and Spanish. At the beginning of each year, teachers provided a general characterization of their students' language in terms of English-Spanish dominance, employing the Student Operational Language Assessment Scale (Duncan & DeAvila, 1976). In late Fall (after becoming more familiar with their students), and again in late Spring, teachers provided a more detailed rating of their students' language skills employing the Oral Language Proficiency Rating Scale (Mace-Matluck, Tunmer, & Domínguez, 1979). Standardized oral language proficiency tests, selected by the district from those allowed by state policy, were administered in the early Fall of each year, the test almost exclusively selected being the Language Assessment Scales (DeAvila & Duncan, 1977). Finally, for a subsample of the target students, audiotaped speech samples were obtained monthly on a rotating schedule in three settings: the classroom, the playground, and the home.

Reading Assessment

Several instruments were used to measure reading achievement. First, detailed information was obtained from two individually administered "performance based" tests assessing both English and Spanish literacy skills.

In the Fall of kindergarten, the Stanford Foundation Skills Test (Calfee & Associates, 1978, 1980; Calfee & Peña, 1978, 1980) was administered to assess pre-reading skills, providing independent measures of visual discrimination, phonetic segmentation, alphabet knowledge, vocabulary knowledge, and narrative comprehension. From the end of first grade on, the Interactive Reading Assessment System

(Calfee & Calfee, 1979, 1981; Calfee, Calfee, & Peña, 1979) was administered during the Spring of each school year to assess reading skills per se. This instrument provided independent measures of skills in decoding, spelling, word meaning fluency in oral reading, and listening and reading comprehension.

As supplemental information, informal reading inventories were administered throughout the school year once the reading of connected text began. Finally, standardized achievement test scores from Spring district-wide administrations were collected yearly in both English and Spanish (though the latter were rarely administered in the study's sites).

Classroom Observation and Teacher Interviews

Project staff conducted monthly observations of the reading instruction in each classroom and interviewed the teachers quarterly about their instructional plans. The observation instrument, the Reading and Mathematics Observation System (Calfee & Calfee, 1976, 1978), documented staffing patterns, grouping and organization, time allocations, the language of instruction, the character of instruction, the materials and procedures used, and the response of the students. The interview instrument, the Reading Teacher Checklist (SEDL, 1978), focused on the teachers' general instructional objectives for reading, as well as the objectives for individual target students. Finally, through the Bilingual Classroom Questionnaire (SEDL, 1979) and its revision as the Inventory of Bilingual Instruction (SEDL, 1981), overall program information was collected through interviews with teachers concerning their daily schedule (as opposed to only their reading instruction periods). Together, these instruments provided a rich characterization of the instructional program for the target students.

Other Data Collected

Oth[illegible] data were co[illegible]d and entered into the data system. For studen[illegible]s, these includ[illegible]y administered cognitive style instruments -- namely, the [illegible]g Familiar Figures Test (Kagan, Rosman, Day, Albert, & Phillips, 1964) and the Children's Embedded Figures Test (Witkin, Oltman, Raskin, & Karp, 1971) -- and a Piagetian cognitive development instrument, the Cartoon Conservation Scales (DeAvila, 1976). For teachers, these included a corresponding set of yearly administered cognitive style instruments -- namely, the adult version of the Matching Familiar Figures Test and the Group Embedded Figures Test (Cons[illegible]ting Psychologists Press, 1971) -- and an instrument designed to ascertain the teachers' background characteristics and language skills, the Survey of Teacher Background and Language Skills (SEDL, 1980). These data have not [illegible]een fully analyzed, some because they appeared not to yield productive information for the purposes of the study under initial analyses, others because of their lower priority in the face of limited resources.

Data Collectors

All direct teacher input data were collected by full-time members of the SEDL research team. This required systematic and frequent visits to the research sites.

Formal classroom observation and the collection of student data were carried out by a data collection team from each of the sites. In most cases, the team consisted of two people who were not otherwise employed. All met the following criteria: resident of the local community, experienced teacher, Hispanic and fluent speaker of English and Spanish, and acceptable to the school district. In all cases, the school district administration provided a list of acceptable and available people who were then screened by the SEDL staff.

Training for the data collectors was extensive and ongoing, with all training conducted by the SEDL research staff, usually at the local site.

Data Management and Reduction

Standard procedures were used in the entry and cleaning of the data, and standard statistical packages (generally SPSS) were then used to obtain descriptive summaries, although a fair number of special-purpose programs had to be written in order to conduct some analyses (most notably, those concerned with the instructional data).

The goal of the data management procedure was two-fold. First, the creation of raw measure data sets that respected the yearly cohort structure of the study (mainly for the purposes of establishing measurement reliability given instrument modifications over the years of data collection). Second, the creation of an integrated data base across cohorts that respected distinctions in grade level (more accurately, instructional year), such that for each individual student a series of measures existed that was in all ways congruent with the series for every other student in the study.

The goal of the data reduction process was to obtain a reliable total score for each subscale of the instruments used in the study, and then, guided by the study's theoretical concepts, to reduce the data structure to a manageable set of relatively independent indices. As such, the analyses to date have treated only the gross features of the data base -- much more detailed analyses treating the data's fine grain have been left for future work.

Data Analysis

In this section, the approaches taken to the analyses of the data base are summarized. First, an overview of the procedure used in the measurement of growth is presented, followed by a description of the techniques employed in the assessment of instruction. Finally, the approach taken to assess the effects of entry skill and instruction on subsequent literacy performance is discussed.

The Measurement of Growth

Given the selection of instruments in oral language and reading, and the summary indices of performance associated with each for each instructional year, the problem of assessing growth for individual students over the semesters in which they may have been tested with a given instrument was next addressed. The solution was derived from the account of growth in the set of component reading skills assessed in the Interactive Reading Assessment System (IRAS), and was then applied to student performance on other instruments as appropriate. As an aid to the explanation of this approach, a brief discussion of the IRAS follows.

IRAS incorporates the developmental dimensions of basal readers for each of the major components of the separable-process model of reading on which it is based: decoding, vocabulary, and comprehension. As an example, in assessing real-word decoding, the IRAS materials were selected using word-frequency lists according to a linear progression in readability (i.e., students were asked to read word lists ordered by the word frequency of their constituent words). To the degree that the basal materials drive student growth in reading, then a year of effective instruction should correspond to a constant amount of progress through the levels of IRAS. That is, the structure of the IRAS materials incorporates a linear component, and therefore, leads to the hypothesis that growth as measured over these material sets should be largely linear.

Accordingly, for each student, performance within a given IRAS subtest over the years tested was assessed by projecting the best-fit regression line through the set of available data points for the given subtest. Such a line summarizes student performance in the relevant task by providing estimates of (a) the intercept, representing the skill level at which the student began schooling (actually, the intercept at first-grade entry was computed rather than at kindergarten entry, since this point seemed to be the modal value of the sample's first systematic instruction in literacy) and (b) the slope, representing the average growth, in terms of IRAS levels, for a single year of instruction.

The degree to which the data actually reflect a significant linear component may be assessed by computing the amount of variability around the individual student's best-fit line. This index is one minus the r-squared value, the latter expressing the squared correlation coefficient between scaled subtest performance and grade level. When converted to a percentage, values of 100% are obtained when the data show no linear component (i.e., the average of the data values is the best estimate of performance for any grade level). At the other extreme of 0% unexplained variance, each of the data values falls precisely on the projected line, and subtest performance is perfectly predictable for any grade level. For values between the two extremes, some linear component is present in the data: As the percentage of unexplained variance decreases, so does the average (linear) prediction error, indicating a relatively larger linear growth component.

For the IRAS growth indices, the average unexplained variance over the English subtests was 19.6% (ranging from 12.9% to 33.9% over the nine component scales); for the Spanish subtests, the average unexplained variance was 26.4% (ranging from 14.8% to 37.9%). These values indicate that, on average, a large degree of growth as measured in the IRAS component subtests can be explained as linear.

Growth in language, as revealed in certain measurement schemes, may not be expected to be linear since the "materials" used to assess language growth may not have been designed in a fashion comparable to the IRAS design. Nonetheless, language growth does show evidence of a linear component, although it is not as substantial as that found in the IRAS indices. For growth in language as assessed by the Language Assessment Scales (LAS), the average unexplained variance for the averaged four multiple choice response scales (Minimal Sound Pairs, Lexical, Phonemes, and Sentence Comprehension) was 23.3% and 32.3% for the English and Spanish indices, respectively; 41.9% and 39.9% for the production rating in English and Spanish, respectively; and 33.2% and 42.9% for the overall level rating in English and Spanish. For the teacher ratings comprising the Oral Language Proficiency Rating Scales (OLPRS), the average percent of unexplained variance was 45.0% and 39.3% for the English and Spanish ratings, respectively. Thus, these sets of language data also reveal linear growth components, though they are not as strongly evidenced as in the IRAS data.

In summary, two points are important to remember when considering the descriptive data based on this analytic procedure. First, the average growth measures presented are based on the best-fit lines projected through each individual student's available data points for a given measure. Second, the fits of these lines to the data are, in general, fairly good, but they do not represent the whole story, as indicated by the varying degrees of unexplained variance.

The Assessment of Instruction

The instruction provided the target sample was documented through two major information sources. The first was through regular classroom observations (approximately five to six per year) employing the Reading and Mathematics Observation System (RAMOS). The second was through regular reading teacher interviews (approximately three per year) employing the Reading Teacher Checklist (Checklists).

The observation instrument consisted of a number of distinct categories of instructional interest (delineated below). Associated with each category was a set of mnemonic codes that detailed the contents of the category (e.g., under the category of Instructional Focus, a large set of codes were used to describe the possible foci, from letter-sound correspondence work to whole word recognition, to text comprehension). For each minute of observation, the observer entered the codes describing the classroom activities under each category, thus providing a minute-by-minute account of the classroom instruction observed. From these sources of instructional data, indices of both the quality and quantity of instruction with respect to a set of instructional dimensions were derived. The particular dimensions of instruction assessed for each individual target student from the observation-based summaries included the following:

Number of Students: the number of students contained in the instructional group.

Classification: the level of the instructor's formal training, ranging from minimal (volunteer) to mid-level (teacher aide) to substantial (substitute teacher, resource teacher, teacher).

Role: the level of formal instruction provided, ranging from minimal (preparation, control, management) to mid-level (facilitation) to substantial (direct instruction).

Subject Matter: the amount of reading generally required by the subject being taught, ranging from minimal (class business, art) to mid-level (science, mathematics) to substantial (reading).

Instructional Focus: the relative explicitness of the instructional emphases and strategies employed in three instructional subcategories:

Letter-Sound Unit: the relative explicitness of the instructional emphasis placed on decoding, ranging from work on isolated units (auditory discrimination, letter recognition, letter-name work) to non-explicit letter-sound pairing (whole word recognition, spelling practice) to explicit letter-sound pairing (letter cluster-sound recognition, letter-sound recognition, spelling pattern recognition).

Word Unit - Meaning: the relative explicitness of the instructional emphasis placed on word meaning, ranging from low (dictionary usage) to mid-level (noun derivative, compound words) to high (antonyms/synonyms, vocabulary enrichment).

Sentence and Text Units - Meaning: the relative explicitness of the instructional emphasis placed on sentence and text meaning, ranging from low (literal facts) to mid-level (story sequence, predicting events) to high (major ideas, making inferences).

Technique: the type of technique in which skills of visual or auditory pattern recognition are presented, as either parts-to-whole or whole-to-parts.

Language of Instruction: the language used in instruction delivery, ranging from all Spanish to alternating usage of English and Spanish to all English.

Materials (Primary and Ancillary): the amount of text contained in the materials used, ranging from minimal (art material, tape recorder) to mid-level (phrase card, chalkboard) to substantial (basal reader, library book).

Activity/Task: the level of formal language demand required by particular activity/tasks in three instructional subcategories:

Non-instructional: the type of instructional activity/task, as either non-instructional (clean-up, wait time) or instructional (all other activity/tasks).

Independent: the level of formal language demand for activity/tasks classified as independent work, ranging from minimal (art activity, copying material) to mid-level (writing from dictation, writing answers) to substantial (test taking, creating writing).

Listening and Responding in Group: the level of formal language demand for activity/tasks classified as listening and responding in groups, ranging from minimal (music activity, playing games) to mid-level (watch-listen, listen-story) to substantial (listen-lecture, discussion-speak).

Attention (Collection Years 1-2): the attention of the instructional group as rated relative to the activity/task required, ranging from low to medium to high.

Number of Nonengaged Students (Collection Years 3-5): the number of students contained in the instructional group who were not engaged in the activity/task being conducted.

Productivity: the rated productivity of the instructional group, ranging from low to medium to high.

Noise: the level of noise as rated relative to the activity/task required, ranging from low to medium to high.

Over the five years of data collection, 1640 observation-based summaries for individual target students were obtained (1293 in English reading and 347 in Spanish reading, the difference reflecting the predominance of English reading offered to these students).

For the interview-based summaries, the particular dimensions of instruction assessed for each individual target student matched those of the RAMOS with the following exceptions, which were not appropriate for the teacher interviews: Subject, Technique, the set of student response indices (Attention, Number of Nonengaged Students, Productivity, and Noise), and transitional activities (Activity/Task: Non-instructional). Two additional categories not found in the observations were included in the interview:

Number of Basals: the number of different basals planned for use in the delivery of the instruction.

Rank: the relative position of the target students' reading group with respect to the following criteria:

> Internal: the relative ranking of the target students' reading group with respect to the other reading groups of the classroom, ranging from low (one of the lowest reading groups) to mid-level (the average reading group) to high (one of the top reading groups).
>
> External: the relative ranking of the target student's reading group with respect to the grade level expectations of the basal reading series employed, ranging from low (below grade level expectations) to mid-level (at grade level expectations) to high (above grade level expectations).

In these interviews, teachers were asked to indicate the general strategies employed in teaching reading to each of the target students, supplying for each, detailed information under the instructional categories of interest (using the same coding scheme employed in the RAMOS), and the relative amounts of time to be devoted to each strategy over the two-week period covered by the interview. Over the five years of data collection, 1943 interview-based summaries were obtained of the instructional plans for providing reading instruction to individual target students (1393 in English reading and 550 in Spanish reading).

Such scaled instructional indices (for both the observation-based and interview-based summaries) represent a set of instructional dimensions with respect to both their quality (reflected in the relative magnitude of the scaled values) and quantity (reflected in the percent of time devoted to each instructional dimension). The use of the term "quality" here does not imply any evaluation of the appropriateness of the instruction, as the skills of the students in a given group may be such that certain types of instruction are obviated. However, this information provides a basis for assessing the kind of instruction received (i.e., its quality and quantity relative to the dimensions defined in this study), and subsequent analyses provided assessments of whether or not instruction so defined influenced the growth of relevant skills of these students.

The findings from the descriptive data based on these summary indices are discussed below. However, these indices were not employed in the integrative analyses. Rather, aggregated indices based on factor analyses (conducted independently for both the observation and interview data, and for English and Spanish reading instruction within each) were used. The summaries derived are described below.

The seven factors identified in the English observation analysis were: (a) engaged text time, an index of reading time where students were engaged with text materials, (b) direct group instruction, an index of direct instruction that was aimed at groups rather than individual students, (c) the quality of formal language, an index of the formal language demands made upon the students, (d) the amount of decoding instruction, (e) student productivity, (f) the use of secondary materials, and (g) the number of students constituting an instructional group.

The seven factors identified in the Spanish observation analyses were: (a) quality of formal language (corresponding to the third English factor derived), (b) direct group instruction (the second English factor), (c) engaged text time (the first English factor), (d) number of students (the last English factor), (e) amount of decoding (the fourth English factor), (f) secondary material usage (no general corresponding English factor) and (g) control (a complex factor also without an English correspondence).

The five factors identified in the analysis of the English interview data were: (a) the amount of comprehension instruction, (b) the quality of formal language, (c) the amount of seatwork, (d) the quality of primary materials, and (e) the amount of group vocabulary instruction.

For the Spanish interview data, the five factors identified in the analysis were: (a) the amount of decoding instruction (the complement of the first English interview factor), (b) the amount of seatwork (corresponding to the third English factor), (c) the quality of primary materials (the fourth English factor), (d) the decoding teacher's classification, which was also associated with the explicitness of the decoding instruction planned (no English correspondence), and (e) the number of students in the instructional group (no English correspondence).

Integrative Analyses

A gross analytic strategy was selected to gain some initial sense of the overall structure of the data base linking precursor skills and instruction to reading achievement. This approach is best viewed as a preliminary analysis of an extraordinarily complex data base -- multiple, yearly student assessments in the domains of cognition, language and reading (in both English and Spanish for the latter two), coupled with extensive yearly instructional data. The primary goals of the analysis were (a) to determine the degree to which the several predictor indices were consistently related to the outcome variables and (b) to evaluate the structural patterns of any such relations.

The primary outcome measures employed in these analyses were the nine summary indices obtained from the nine subtests found in the IRAS, analyzing English and Spanish performance separately. For convenience in interpretation, these indices may be grouped into three major categories: (a) oral language (Vocabulary Definition, Narrative Listening Comprehension, and Expository Listening Comprehension), (b) Word/Sentence Decoding (Vocabulary Decoding, Synthetic-word Decoding, Synthetic-word Spelling, and Sentence Reading), and (c) Reading Comprehension (Narrative Reading Comprehension and Expository Reading Comprehension).

For each year that a student was tested, a deviation was computed between each of the student's IRAS measures and the aggregate growth track index summarizing average performance during that year. The IRAS deviations for each year were then submitted to a regression analysis in which the predictors included (a) precursor indices (oral language classification at entry to school and previous year's performance on the corresponding IRAS measure) and (b) instructional dimensions (the nominal reading program, the observation-based and interview-based indices of instruction, attendance, and site at which the attended school [illegible]ted).

The major disadvantage of this approach can be easily stated: Because each year is considered in isolation from the others, there is a loss of information about the longitudinal character of changes in reading achievement. Likewise, there is a loss of information about the configurational patterns relating changes in instruction to changes in achievement. Other analytic procedures are possible, but they require more resources than were available for these analyses.

Preliminary to the regression analyses proper, the correlations between the set of predictors and the set of outcome measures were carefully examined, as were those among the set of predictor variables. The derivation of each of the predictor variables is briefly described below.

Precursors

The first precursor, oral language level on entry to kindergarten, was determined for both English and Spanish as a two-level category (median split) based on teacher ratings. As it happened in this sample, oral language competence was virtually independent on entry to kindergarten. Analyses of variance conducted for each of the IRAS longitudinal measures with English and Spanish language ratings as the independent factors showed that the English IRAS measures were generally affected by the English language rating, and the Spanish IRAS measures were generally affected by the Spanish language rating with little evidence of systematic interactions between the two. Accordingly, in all of the regression analyses, oral language as a precursor was simply represented by the corresponding language rating around the median split.

A student's achievement level at the end of a given school year is generally related to performance at the beginning of the year. Accordingly, an index of previous performance was included as a precursor. From second grade on, the corresponding IRAS deviation served as the index. For first grade achievement, the Alphabet Knowledge subtest from the Stanford Foundation Skills Test (SFST) was employed. For English, awareness of the letter names is known to be correlated with later reading achievement, for reasons that are not entirely clear. The distribution of scores on this subtest was bimodal in this sample, replicating earlier findings, and so this precursor was reduced to a dichotomous contrast.

Instruction

Two indices of a student's status in a bilingual reading program were employed in the regression analyses. The first was the total number of years of assignment to Spanish reading instruction of any sort (ranging from 0 to 5), and the second was a dichotomous variable indicating whether the student was assigned to Spanish reading instruction during the particular instructional year under analysis.

Specific instructional dimensions were based on the seven RAMOS and the five Checklist factor scores. English instructional summaries were used in the regression analyses for English IRAS deviations and Spanish summaries for Spanish IRAS deviations.

Attendance data were provided every year by the districts for each student, and the percentage of days attended during the given instructional year was used as the index of school attendance.

Finally, since the study was conducted at six different districts, a set of orthogonal contrasts were introduced as the last step in the regression analyses in order to assess any other between-site effects that were not included as part of the other influences (i.e., the precursor and instruction indices).

Summary

In summary, the data analyses allowed various perspectives on the data base, and results from these will be summarized below for both English and Spanish indices as follows. First, the linear growth of students in oral language skills as assessed through various instruments will discussed. Second, descriptive data on the reading instruction received by these students will be summarized. Third, for reading skill, the following will be discussed: (a) descriptive data on entering "reading readiness" skills, (b) the linear growth of students with respect to component reading skills, (c) the relationships between entering reading readiness skills and subsequent linear growth in the various components of reading, and (d) the relationships between linear growth in Spanish reading and linear growth in English reading for the various reading component skills. Finally, the relationships of entering skills and instruction with subsequent reading achievement will be discussed as revealed in the integrative analyses.

ORAL LANGUAGE

The students in the bilingual sample were deemed by their schools to be Limited English Proficient, as determined by their performance on a standardized test of oral proficiency given in the Fall of their kindergarten year, and they were therefore enrolled in bilingual kindergarten or first grade classrooms when they entered the study.

As noted above, three types of language measures were used in the study: (a) an oral language proficiency test, (b) teacher ratings, and (c) audiotaped interactions -- language samples. Analyses of the data from these measures revealed the following:

- o.. The students in the sample, on entry into school, varied considerably in their degree of bilingualism.

- o.. The students, generally, made considerable progress in acquiring skill in English; less growth was observed in their performance in Spanish.

- o.. Site differences were apparent in the students' facility in Spanish and in English on entry and in their subsequent growth in each language.

- o.. Site differences were also observed in the patterns of language choice, both at entry and over time.

- o.. The student's oral language proficiency varied, in both languages, as a function of the type of measure used (oral language proficiency test versus teacher ratings) as well as by the type of task within a given measure (story retelling versus discrete items that required short-answer responses).

- o.. When compared to teacher ratings, the oral language proficiency test used appeared to underestimate the students' ability in both languages at entry and, at higher grades, to overestimate their English abilities and underestimate their skills in Spanish.

A number of critically important instructional issues surround language assessment. Primary among these is the question of adequate and accurate assessment of the oral language abilities of young children. Objective measures, such as the currently available standardized oral language proficiency tests, have been widely criticized. The widespread dissatisfaction with these measures arises from the belief that these tests do not reflect the totality of the language resources that children possess, nor do they adequately predict children's ability to perform in the school setting. Further dissatisfaction arises from the concern that formal testing of young children's language may in fact be measuring many things other than language (e.g., general readiness for school, knowledge of test taking). Subjective measures, such as teachers' ratings, have been maligned by some who point to the "human element" that comes into play with such procedures. Natural, or free speech, samples avoid some the potential pitfalls of other types of measures, but they, too, have their limitations.

ISSUE: Valid Language Assessment

The research staff, fully aware of the limitations of the various kinds of measures and of the hazards involved in oral language assessment (given the state of present knowledge about what constitutes oral language proficiency and how to assess it), employed multiple measures in an attempt to obtain a reasonably accurate index of each student's oral language abilities and patterns of language choice over time. Analyses of the oral language data strongly suggest that none of the existing measures by themselves provide adequate information on which to base educational decisions. Use of a variety of types of measures and procedures can, we believe, provide a reasonably accurate index of the student's oral language abilities. However, this process is time consuming and requires skill and expertise that often is not readily available within most school districts.

IMPLICATIONS: Given that results of oral language assessment figures prominently in a number of educational decisions regarding schooling practices for language minority children (e.g., identification, program placement, termination of special services), further research is urgently needed to determine not only effective but practical means for assessing the oral language proficiency of young children.

ISSUE: Language and Instructional Program

When examining language as a precursor skill for reading achievement, additional instructional issues emerge. First, to what extent does the child's language at the time of entry into school determine program placement? By legislative mandate, all children in Texas from non-English language backgrounds who, at entry into school, score at or below a predetermined cutoff score in English on the district-selected oral language proficiency test are placed in a bilingual education program, which implies some use of the home language (e.g., Spanish) for instruction for some given period of time. Thus, the issue here is not one of access to the program but rather the accuracy and adequacy of the information on which placement decisions are made.

IMPLICATIONS: School districts should be strongly encouraged to use multiple kinds of data from various sources in arriving at decisions about the placement and instructional treatment of language minority children. A formal language measure can provide one kind of information; professional judgement of school personnel about the student's language characteristics in both formal and informal settings in the school environment can provide another; the student's academic performance can reveal further information; and home language surveys and educational histories contribute additional important data. Hence, oral language proficiency test scores should not be the only (nor even the primary) source of information on which decisions are made.

Second, to what extent does the child's language at the time of entry into school determine the actual instructional program delivered? The teacher's perceptions of the child's language abilities and instructional needs determine to a large extent the instructional treatment delivered to the student. Therefore, in bilingual classrooms, use of the home language for instruction for a given child or group of children will vary, both as a medium of instruction and for support within the classroom environment. Underlying transitional bilingual education programs is the belief that reading is a single process and that having learned to read in one language, reading in another known language is a matter of transferring and extending one's knowledge and skills. It is also generally believed that bilingual children learn to read more easily and more efficiently when their initial reading instruction is provided in their stronger language. Therefore, transitional bilingual programs may provide initial reading instruction in Spanish for children who are clearly Spanish dominant and are limited in their English skills at the point that formal reading instruction is begun (usually in first grade). In the present study such instruction was provided for approximately one-third of the students. While all of the students in the sample were deemed by their schools to be Limited English Proficient on entry into school and were enrolled in bilingual classrooms when they entered the study (as kindergarteners for the most part), subsequent placement and instructional decisions resulted in initial reading instruction in English for the majority of the students.

Length of stay in the program is also determined to a large extent by the teacher's perceptions of a student's readiness to perform in an all-English classroom, as well as by prior instructional treatment and the student's progress in acquiring English. In the present study, some students were transferred to a regular mainstream program at the end of their kindergarten year, presumably because they had either made rapid progress in acquiring English and were therefore no longer considered Limited English Proficient or because their English skills had been underestimated in their entry language assessment. Other students in the sample scored low in both languages (or were perceived by their teachers to have attained less than adequate oral language development in either language), and it was presumed that English reading instruction would be as appropriate for these children as would Spanish reading instruction. Yet other students who remained in bilingual classrooms in first grade and received initial reading instruction in English may have gained sufficient skills in English to begin such instruction but were deemed in need of support in the home language in other curriculum areas. Thus, contrary to popular belief, not all children enrolled in bilingual classrooms receive reading instruction in their non-English home language. Of the students in the present study who did receive Spanish reading instruction, most remained in such reading programs for at least two years.

IMPLICATIONS: Given that teachers' perceptions determine to a large extent the instructional treatment delivered to children, teachers should be knowledgeable about and have significant input into the oral language assessment process, particularly as it relates to entry/exit decisions.

Teacher training in the area of oral language assessment needs to be strengthened. Such training should equip teachers to be astute observers; they need training not only in how to observe language behavior, but also in the content of what to observe (e.g., social, personal, and cognitive aspects of oral language proficiency as well as the linguistic dimensions).

Research is needed that:

- o.. focuses on how to train teachers to be better observers of language performance, including the content of what is to be observed.

- o.. examines criteria teachers use in making placement (grouping) decisions for reading instruction within a class and for instructional treatment decisions relative to each of the instructional groups.

- o.. examines, on the state level, how teachers make decisions about bilingual and English as a Second Language students and the extent to which those decisions are congruent with state and local policy; and identifies ways in which rules, regulations, and teachers' decisions may converge in order to arrive at consensus among practitioners and regulatory agencies.

ISSUE: Language Development and Reading Acquisition

To what extent does the child's language development at the time of entry into school affect subsequent reading achievement? The literature is replete with studies that have shown a moderate-to-strong relationship between oral language development and reading achievement. Knowledge of the language being read is at the heart of the reading process. Reading is a derived skill in that it builds upon oral language and requires the translation from writing to a form of language from which the reader already is able to derive meaning. To learn to read, children must bring their knowledge of the spoken language to bear upon the written language. A well-developed system of oral language assumes a functional vocabulary and the ability to discover the structure and meaning underlying spoken utterances. It also assumes a rudimentary ability to reflect upon language that allows children to discover the properties of spoken language that are central to the correspondence between its written and spoken forms (e.g., awareness of relationships among words in text, as well as among higher-order structural units such as clauses and sentences). Children who do not have a well developed understanding of the communicative process at entry into school often experience difficulties in learning to read and therefore fall below the school's expectations in their academic progress. In the present study, the oral language skills of approximately 25% of the students in the sample were, at entry, rated relatively low in both languages by their teachers.

IMPLICATIONS: School-based preschool programs, along with parent involvement components of school programs, have gained support as a means of enhancing the language development of young children. With adequate attention to staff development, instructional focus, monitoring, and funding, such programs could significantly advance the language development of "high risk" youngsters and should therefore be made more widely available to low-income language minority students.

Research is needed that examines the effects of preschool programs on the language development of language minority children and related effects on subsequent reading achievement.

INSTRUCTION

A coordinated system of classroom observations and teacher interviews provided rich and extensive data on the instructional program each target child received over the course of the study. The major findings concerning the dimensions of instruction assessed in the study are summarized below.

Of the 250 bilingual students in the sample, 70 students (28%) began reading instruction in Spanish. Of those, some received Spanish reading instruction for one year before being transferred to English reading; others remained in Spanish reading for two, three, or four years, with most being transferred to English reading by the end of third grade.

Analysis of the instructional data suggests that the teachers generally implemented the instruction they had planned, as indicated by a close match between the instruction observed and what teachers said that they were going to do during a given period. Instruction in Spanish and English was similar in terms of the instructional dimensions assessed in the study.

In general, the instruction delivered may be characterized as follows:

o.. The size of the instructional group for reading ranged from about 13 in the early grades to about 15 in the later grades.

o.. The teacher (as opposed to an aide or some other "instructor") was associated with the target students about two-thirds of the time.

o.. The role played by the teacher tended to be one of facilitation (rather than direct instruction) in the aggregate.

o.. The language of instruction tended to be English during the English reading classes and Spanish during the Spanish reading classes, but with some English used during the Spanish reading period.

o.. About half of the instruction time devoted to reading instruction during the first and second grades was focused on decoding; this fell to about 30% by fourth grade.

This instruction tended to involve non-explicit letter-sound pairings at each grade level (e.g., children were shown a word on a flash card and were asked, "What does this word say?" or were asked, "What is the name of this letter?"); little explicit letter-sound work was observed (i.e., direct instruction in or practice on tasks that required the child to overtly focus on a specific isolated linguistic element and relate it to its graphic representation).

o.. The amount of time spent on developing word meaning was small.

o.. About 30% of the instructional time was on instruction in the meaning of sentences and texts (comprehension) in the first three years, with a slight increase in the fourth year.

The quality of this instruction was fairly stable across years, and was generally non-explicit (e.g., favoring a focus on literal facts over making inferences).

o.. Independent work accounted for about half of the instructional time during the first two years, dropping to about 35% in the following years.

o.. The level of formal language demand (i.e., the extent to which the activity or task required interaction with connected instructional text either oral or written) observed in both independent and group work was low, starting at a relatively low level in the first year, and increasing to mid level by the last two years.

o.. The primary materials used in instruction tended to be basal readers accompanied by workbooks, worksheets, and chalkboard/paper/pencil activities.

o.. The number of nonengaged students was low; productivity was rated medium each year, while noise tended to be low.

In summary, for both the English and Spanish observational data, instruction was largely conducted by a teacher acting in a role of facilitation. Over instructional years, an increased reliance on group work over independent work was seen. Much of the early grade work was focused on decoding, declining in the latter instructional years. The quality of this instruction tended to be non-explicit as little instruction dealing specifically with letter-sound correspondences was seen. Little instruction in vocabulary was observed, although that which was observed tended to be quite explicit. Finally, instruction in sentence/text meaning complemented the time devoted to decoding, showing a small increase over instructional years; like decoding, this instruction was generally non-explicit, favoring a focus on the literal facts contained in the text material as opposed to a synthesis of its major elements.

One of the most important findings revealed in these data concern instructional explicitness, the degree to which instruction is specific in its detail. For decoding, such explicitness is the degree to which the correspondences between letter(s) and sound(s) are isolated for the student. For comprehension, explicitness represents the degree to which text structure is isolated. Note that for both decoding and comprehension instruction, such explicitness was, in the aggregate, low. Thus, the task of determining the underlying relationships expressed in the instruction offered were left to the student with relatively little instructor assistance.

The summary data on which the above descriptions were based were subjected to factor analysis in order to reduce the number of instructional indices. The observational and interview data were treated separately, as well as those representing English and Spanish instruction. For both the English and Spanish observational data sets, seven factors were derived. The seven factors identified in the English analysis (in their order of strength) included the following:

o.. Engaged Text Time, an index of reading time where students were engaged with text materials.

o.. Direct Group Instruction, an index of direct instruction delivered by an instructor and which was aimed at groups of students, rather than individuals.

o.. Quality of Formal Language, a measure of the formal language demands made upon the students.

o.. Amount of Decoding, a measure of the relative amount of time devoted to instruction in decoding.

o.. Productivity, an index of the conditions promoting high individual student productivity.

o.. Secondary Materials, a measure of the relative usage of secondary materials.

o.. Number of Students, an index of the number of students constituting an instructional group.

The factor solution derived from the Spanish observational data was similar to that derived from the English observational data, with five of the seven factors containing many of the same component variables. The seven factors identified in the Spanish analysis included:

o.. Quality of Formal Language (corresponding to the third English factor derived).

o.. Direct Group Instruction (the second English factor).

o.. Engaged Text Time (the first English factor).

o.. Number of Students (the last English factor).

o.. Amount of Decoding (the fourth English factor).

o.. Secondary Materials, an index of both the quality and quantity of secondary material usage (only tangentially related to the sixth English factor).

o.. Control, a complex factor without an English correspondence, which is essentially an index of the number of management interruptions.

Although these factors are interesting in their own right, their main function was served in subsequent analyses where they represented instructional dimensions in order to ascertain the degree to which instruction was related to student achievement in reading. The discussion of these relationships will be deferred to a latter section of this report.

The Spanish and English teacher interview data were also analyzed separately, and in both analyses, five factors were derived. The five factors identified in the English analysis included the following:

o.. Amount of Comprehension, a measure of the relative amount of planned time to be devoted to instruction in comprehension.

o.. Quality of Formal Language, an index of the formal language demands required by the planned instruction

o.. Seatwork, an index of the relative amount of time to be devoted to independent seatwork as opposed to group work.

o.. Primary Materials, an index of the planned usage of primary materials.

o.. Group Vocabulary, an index of the relative amount of time to be devoted to group instruction in the meaning of words.

The factor solution derived from the Spanish interview data set was similar to the English interview solution, and included the following five factors:

o.. Amount of Decoding (a compliment of the first English factor).

o.. Seatwork (corresponding to the third English factor).

o.. Primary Materials, an index of the quality of the primary materials to be employed.

o.. Decoding Teacher Classification, a factor defining the relative educational training of the teacher expected to deliver decoding instruction (which was associated with the explicitness of such planned instruction).

- Number of Students, an index of the relative instructional group size.

Thus, these factors represent indices of the instruction offered the study's target children (with respect to both quality and quantity) along a number of dimensions that, by hypothesis, should be critical for the development of reading skills in the early grades.

READING

A primary purpose of the study was the investigation of patterns of growth in reading achievement. The study employed multiple measures for assessing each of the major components of skilled reading (vocabulary knowledge, decoding, and text comprehension). For the bilingual sample, such growth was monitored in both English and Spanish.

Pre-reading Skills Development

Underlying general skills that are critical to acquiring new knowledge and skills (e.g., attention, memory, verbal fluency, effects of previous learning) are a set of independent component skills that are intimately related to the acquisition of reading. These include decoding, word meaning or vocabulary, sentence and paragraph comprehension, and text comprehension. Assessment of the students' pre-reading skills at entry into the study revealed the following:

- The overall sample of students came to school with sufficient skills to begin literacy acquisition -- they did not appear to be academically disadvantaged.

- Approximately one-half of the sample of students came to school knowing the letter names of the English alphabet, which has been found to be a good predictor of early English literacy exposure.

- Knowledge of the Spanish alphabet was negligible, but expected, given its different treatment in the language and culture.

- Sight-word recognition was minimal in both languages, but higher in English than in Spanish.

- Visual matching skills were already highly developed.

- Auditory segmentation skills could readily be acquired with familiar words by most of the students, with higher performance with English words than with Spanish words. The transfer of this skill to novel items was difficult for some.

1082

o.. Vocabulary knowledge was high, with slightly greater strengths in English.

o.. The formal dimensions of schooling and text (as measured by listening comprehension) appeared to be new to many.

o.. Visual matching tasks and the metalinguistic task of auditory-phonetic segmentation possessed a degree of transferability between the two languages, while the linguistic tasks tapping vocabulary knowledge and comprehension were independent across (but not within) the two languages; alphabet knowledge and sight-word recognition tended to be related across the two languages.

Reading Achievement

The Interactive Reading Assessment System - IRAS was used to measure the components of skilled reading. Standardized reading achievement scores were collected yearly wherever available. Summaries of the data obtained from these measures are presented below.

IRAS-English

o.. For the overall sample, the students entered first grade with English oral language skills that exceeded the expectations of the growth track model but grew in accord with the model predictions; thus oral language skills were above grade-level expectations throughout the primary grades.

o.. The decoding skills of the students were minimal at first-grade entry, and they showed subsequent growth that was above grade-level expectations (progress in spelling, however, was slow); thus decoding (of isolated words) was, like oral language skills, above expectations throughout the primary grades.

o.. Decoding fluency may have presented problems in reading connected text as by second-grade exit, the average student had a reading rate of less than two syllables per second.

o.. Reading comprehension was about half a grade level below expectation at entry and showed growth slightly above the expected rate; thus reading comprehension was found to be slightly below grade level expectations throughout the primary grades. At fourth grade exit, the overall sample was projected to be within a half grade of that expected by the growth track model.

o.. Students who came to school with relatively lower English skills showed greater growth in English oral language capacities, and they showed a convergence in such skill in late fourth grade with those students who entered with higher English skills. However, the high English entry students were better able to profit from decoding instruction in that their initial advantage in decoding continued to expand.

o.. A significant relation was found between entry level English skills and narrative reading comprehension in English. The high English entry group entered first grade with reading comprehension skills at the expected level, while the low English entry group was about two levels below expectation. Growth rates did not differ; each proceeded at the rate of about a year of growth for a year of instruction. At fourth grade exit, the high English entry group was projected to be about a half grade level above the growth track model expectations, and the low English group was projected to be about one grade level below.

o.. Students with relatively higher Spanish oral language skills at entry into kindergarten had growth rates in English reading comprehension that exceeded those of students with relatively lower entry Spanish oral skills. This suggests that although the development of English listening comprehension did not differ for these groups, relatively higher skills in Spanish at school entry promote the growth of English reading comprehension.

IRAS-Spanish

o.. In the overall sample, the students entered first grade with Spanish oral language skills that exceeded the expectations of the growth track model but grew at half the expected rate; thus oral language skills were above grade level expectation at entry, but were projected to fall below grade-level expectations during the primary grades.

o.. The decoding skills of the students were minimal at first-grade entry and showed subsequent growth that was slightly above grade-level expectations (as in English); thus decoding (of isolated words) was above expectations throughout the primary grades. Progress in spelling, however, was slow.

o.. As in English, the data suggest that decoding fluency may have presented some difficulty in reading connected text.

o.. Reading comprehension was a grade level below expectation at entry and showed growth that was only half the expected rate; thus, reading comprehension was found to be substantially below grade-level expectation throughout the primary grades.

o.. Entry level skills in Spanish were related to reading performance in Spanish. The low Spanish entry students

1084

entered first grade with less skill than the high Spanish group in the areas of formal language and decoding, but subsequent growth did not differ. For reading comprehension, however, the two groups began with the same low-level skills, but, given the greater formal language and decoding skills of the high Spanish entry group, their growth in reading comprehension was able to proceed at a greater rate. This rate, however, was substantially below that expected from the growth track model, and the data suggest that the major difficulty for these students was not decoding skill, but rather, skill in dealing with the formal language aspects of text.

Relations within and between IRAS Measures (English and Spanish)

o.. Within both English and Spanish, the relationships found between the component scales can be summarized as follows. The highest relationships were generally between the component scales within the three major skill areas assessed (formal language, decoding, and reading). The correlations between these skill areas were strongest for decoding and reading, somewhat weaker between formal language and reading, and weakest between decoding and formal language. Thus, the general correlational patterns suggest that decoding and formal language skills are relatively independent, with both needed for growth in reading comprehension.

o.. The correlational pattern between the English and Spanish indices was as follows: skill in decoding, and to a lesser degree, in reading, was related across the two languages, while formal language skills (as expected) was generally unrelated. There was, however, a general trend for stronger relationships between a given English task across the set of Spanish tasks when compared to those relationships for the same given Spanish task across the set of English tasks -- this suggests that literacy development in English may be more readily transferable to Spanish than from Spanish to English.

Relationship between Pre-reading and Reading Achievement Measures

The correlational patterns between the pre-reading measures and the reading measure (IRAS) was as follows:

o.. Knowledge of the letter names of the English alphabet at kindergarten was found not only to be generally related to English literacy skill at first-grade entry, but also to subsequent growth in decoding and reading acquisition. Knowledge of the letter names in Spanish, however, did not carry such widespread predictive power for Spanish literacy development, neither for entry skill nor for subsequent growth.

- o.. Kindergarten entry skill in decoding and oral language was related to such skill at first-grade entry, within both English and Spanish; but for English, some of these entry skills were further related to subsequent English literacy growth (segmentation to decoding, and oral language to reading growth).

- o.. For the cross-language correlations, few significant relationships were found.

Standardized Reading Achievement Tests

Generally beginning in first grade, standardized achievement tests were administered to all students in the Texas sites in the Spring of each year. Three different standardized tests were used over the course of the data collection phase: the California Achievement Test, the Comprehensive Test of Basic Skills, and the Iowa Test of Basic Skills. Standardized achievement tests in Spanish were not administered systematically, nor to any great extent, by any of the schools in the study. Performance of the students on the standardized reading achievement tests in English may be summarized as follows:

- o.. Performance in English indicated that the students in the overall sample entered first grade just slightly below grade-level expectations and showed growth which was also slightly below expectation. By fourth grade exit, the sample was projected to be a full grade level behind.

- o.. A significant relation was found between entry level English skills and reading performance in English, as measured by standardized reading achievement tests in English. The low English group began first grade just below grade level expectation, with subsequent growth that gave about three-quarters of a grade-level improvement for each year of instruction. The high English group began first grade slightly above grade-level expectations and grew at a rate that was slightly below expectations. Thus, at fourth-grade exit, the high English entry group was projected to be about a half grade level behind, while the low English entry group was projected to be slightly more than one grade level below grade norms.

ISSUE: Pre-reading Skills Development and Reading Achievement

To what extent does the child's pre-reading skills development at entry affect subsequent reading achievement? As noted above, children's knowledge about literacy at entry into school has an impact on their reading achievement both in the early stages of literacy acquisition and in later reading achievement. An important question for educators is, "Can instruction change the relative level of attainment in literacy that is predicted by individual differences between children in their knowledge about literacy on entry into school?" A number of studies have shown that differential progress in the acquisition of literacy is related to the quality of instruction delivered to children. The problem remains, however. Children who are well prepared at entry to take advantage of what the school has to offer progress at the rate of approximately a year of growth for a year of instruction; children less prepared often get off to a slow start, and even if they progress at the rate of a year of growth for a year of instruction, they still lag behind their more advantaged peers as they progress in school. Furthermore, these children often get locked into an instructional track in which the instruction is the same as for other students, only at a slower pace; the students' own expectations, as well as those of others for them, decline; and the range of instruction delivered is such that these students have limited opportunities to learn more than a narrow range of the skills and content needed to become fluent readers.

IMPLICATIONS: A challenge for the schools is to find means for helping the less academically advantaged children become better prepared to benefit from instruction and for accelerating their growth in the early years so that they can keep pace with the general school population of their own age. Well-designed preschool programs could expand the knowledge and experience base needed for subsequent academic learning. Further, the whole concept of ability grouping for instruction and instructional "tracking" needs to be subjected to close scrutiny. Such organizational procedures may not be in the best interest of low achieving students.

ISSUE: Rate/Pattern of Language and Reading Growth

To what extent does growth and development of oral language following school entry contribute to reading achievement? Children from a non-English language background who enter school with limited English-speaking skills face the task of acquiring mastery of the grammar of a new language, but the problem goes far beyond that. It is becoming increasingly clear that the classroom is a unique communication setting. As yet we know only little about the nature of the linguistic demands that are placed on students during the very critical years of their schooling, and about how bilingual children acquire competence in using languages for both academic and social purposes. Most children by the age of five have achieved control of the basic structure of their native language and of most of the complexities of conversational interaction. However, upon entering school, children confront a new speech environment with different linguistic requirements for accomplishing their informational and social goals. They must, therefore, acquire other dimensions of language not yet developed by many children in their native language at school entry. Two examples may suffice to demonstrate this point and to define some of the features of what we have termed "formal" or "school-related" language.

One example has to do with the interactional rules (or interactional requirements) of the classroom. Although there appears to be no significant "transition" involving grammar, phonology, or even vocabulary in a given language between that required for communication in the home and in the school, differences do occur in the nature of the interactions that take place in the two environments. First, there are differences in the relative frequency of types of interaction. For example, the three-part structure of the question-answer sequence (teacher-pupil-teacher), occurs with far greater frequency in the classroom, as do pseudo-questions where the asker already knows the answer (e.g., Teacher: This is a triangle. Billy, what is this?).

Second, the role or status of the participants differ in the two settings. In interpersonal communication, typical of the home, the child shares the responsibility for initiating topics; there is a two-way flow of new information, and meaning is often supported by shared knowledge of the event, as well as by contextual cues from the situation in which the exchange takes place. Quite the reverse is true in the classroom. Most often the teacher is the topic initiator and assumes the authority role -- the student is the recipient; the flow of new information is a one-way event (knowledge of the event may not be shared by the child); and contextual cues are greatly reduced.

Third, the conversational structure of classroom talk differs, partly due to the pedagogic motivation that underlies much of the talk at school but also due to the special requirement of maintenance of order in conversation involving large numbers of participants. Thus, children must learn a set of discourse rules that are required for that particular setting (e.g., how to successfully engage the teacher and others to acquire the necessary input for learning; when and under what conditions a turn can be successfully negotiated; how to deal successfully with a specified topic). At entry into school, student competence varies in these special aspects of communication associated with classroom activities. Success in school is dependent upon adequate knowledge of the rules of classroom discourse rules, and for many children this is a major learning task in their early years of schooling.

A second characteristic of "school-related" language has to do with the use and interpretation of language in different contexts. In recent years it has become increasingly clear that users of language acquire skill in both the natural and formal domains of speech and thought. Children, having been raised in the informal and intimate language of the home, come to school with linguistic skills characteristic of natural language, but it is formal language that is used in oral discourse in the classroom and in the textbooks of school. It is this form of language students have to acquire in order to make academic progress.

A number of scholars studying the relationship between language and thought have drawn a distinction between the use and interpretation of language used in face-to-face communication and language that is used autonomously. In the former, the language used is supported by contextual and paralinguistic cues and is therefore less dependent on the specific linguistic forms used for its interpretation than it is on the expectation and perception of the speaker's intentions and the salient features of the context. In contrast, language and thought that moves beyond the bounds of meaningful interpersonal context (i.e., formal language) makes different demands on the individual and requires the user to focus on the linguistic forms themselves for meaning, since meaning is autonomously represented and contextual support is greatly reduced. The linguistic message must, therefore, be elaborated precisely and explicitly, whether in the oral or written form.

To a considerable extent, formal education is concerned with teaching the child to process and to produce those varieties of spoken and written language in which meaning is autonomously represented. Growth in language equips the child to use language symbolically to represent remote, imaginary, or even hypothetical events and experiences. In the acquisition of literacy and the spoken form of formal language, children learn to assign meaning to the linguistic forms per se and are made conscious of the process by which language can be controlled and manipulated to gain knowledge and to apply that knowledge in a variety of academic and social contexts. Learning to deal with language in this manner is essential for success in reading, yet it is a difficult process for many children, since it requires learning to view and to use language in a new and expanded way.

As noted above, the oral English skills development of the overall sample of students in the study (as assessed through listening comprehension) proceeded at a rate that exceeded the expectations of the growth track model. The students made considerable progress not only in learning English grammar but also in acquiring the dimensions of English proficiency referred to above as "school-related" language (as reflected in teacher ratings and ratings of audiotaped language samples; also Klee, 1984). Their oral Spanish skills, in contrast, proceeded at only half the expected rate and were projected to fall below grade-level expectations during the primary grades. The lesser growth in oral Spanish skills can be explained in part by the fact that only about one third of the sample received varying amounts of literacy instruction in Spanish in the early grades, consequently a disproportionate amount of the instruction for the entire group was in English over the course of the study. Nonetheless, mastery of a second language beyond that required for interpersonal communication takes time. Even with considerable emphasis on English reading instruction in the classrooms in the study, the oral English skills of the Low English entry group, while showing greater growth rates than that of the High English entry group, did not match those of their relatively more English proficient peers until late fourth grade. This finding lends support to other research that suggests that it takes young children several years, on the average, to approach grade norms in cognitive/academic skills in their second or weaker language (Cummins, 1983).

While growth in oral English skills proceeded at a rate that exceeded expectations, entry oral language skills, nonetheless, had a pervasive effect on various aspects of reading achievement. These skills were associated with entry level skills in decoding in both English and Spanish; students who entered with relatively higher oral skills tended to have relatively greater skill in decoding than did students with relatively lower oral skills in a given language, and for English, rate of growth was similarly affected. However, the students in the sample, on the average, acquired the ability to decode words in isolation in both languages at rates that exceeded grade level expectation and were found to be above grade level expectation in decoding throughout the primary grades. Such performance is typically observed in reading in a second language. However, the rate at which children are able to decode and process linguistic units (decoding fluency) affects reading comprehension. The reading rate of the students in the study was slow in both languages (viz., less than two syllables per second by the end of second grade). Potential factors contributing to a slow reading rate include lack of automaticity of decoding skills, inadequate word analysis skills, less than adequate development of oral proficiency, limited knowledge of text structure, and insufficient use of enabling text processing strategies (e.g., use of context cues, noting overriding themes, adjusting reading rate to the purpose of the task).

In the aggregate, English reading comprehension, as assessed by the performance-based measure (IRAS), showed growth that was slightly above the expected rate. However, the students' entry level reading comprehension skills were such that, progressing at the rate of a year of growth for a year of instruction, their reading comprehension remained slightly below grade level expectations throughout the primary grades.

Oral language entry skills were shown to be substantially related to performance in reading comprehension. Students who entered with relatively high English oral skills also entered with better developed reading comprehension skills and, while growth rates did not differ for the two groups, the advantage of the high English group at entry resulted in a consistently higher level of achievement in reading comprehension for this group. On the other hand, children who entered with relatively higher Spanish oral skills had growth rates in English reading comprehension that exceeded that of students who entered with less well developed skills in Spanish. These findings suggest that children who come to school with well developed oral language skills in either or both languages have an advantage in learning to read connected text, as higher oral English entry skills were associated with higher entry reading comprehension skills, and higher oral Spanish skills were associated with greater growth in English reading comprehension (but not with entry reading comprehension level).

Finally, for standardized tests of English reading, higher entry oral English skills were associated with both higher English reading comprehension entry skills and higher growth rates. Spanish oral skills had no influence on either entry level comprehension skill or growth rate on standardized English reading achievement tests. It is generally expected that effective instruction will produce at least a year of improvement for a year of instruction. For high English entry students, who for the most part received their initial reading instruction in English, growth in English reading approached the expected rate. For low English entry students, rate of growth in English reading was about three-quarters of a grade-level improvement for a year of instruction. This can be explained in part by the fact that a portion of this group received reading instruction in Spanish for one or more semesters (usually two or more years) before being transferred to English instruction.

IMPLICATIONS: For the students in the study, present schooling practices are resulting in reading achievement on performance-based tests projected to be within a grade level of expectation at the end of fourth grade, with students who entered with relatively higher English skills projected to be slightly above grade level expectations. On standardized reading achievement tests the picture is somewhat different. The overall sample is projected to be about a grade level below expected norms, with the high English entry group approaching grade level norms by the end of fourth grade.

Several studies (Doebler & Mardis, 1980-81; González, 1977; Leyba, 1978; Rosier & Farella, 1976; San Diego City Schools, 1982; Troike, 1981; Willig, 1985) suggest that the full benefits of initial reading instruction in the home language often are not apparent until students who have received such instruction are in the later elementary grades. A follow-up study of students in this study could provide greater insights into the long-term effects of bilingual instruction.

Further, some students in the study were deemed by their teachers to have had low oral language skills in both their home language and English on entry into school. Since entry oral language skills were found to be associated with reading performance in important ways, research that can assist schools in working effectively with such students appears to be warranted.

Similarly, effective means for increasing the reading rate (decoding fluency) of children from non-English language backgrounds need to be identified and communicated to teachers, as less than adequate reading rates may be impeding growth in reading connected text for such children.

ISSUE: Transfer of Skills Across Languages

To what extent do knowledge and skills gained in one language transfer to similar tasks in another known language? Underlying transitional bilingual education is the premise that reading skills gained in initial instruction in the home language can be transferred to reading in English and that children, having learned to read successfully in their home language, can be taught to read at the same level relatively easily in English once oral English skills have reached an acceptable level of proficiency. This assumes that transfer of learning will occur when certain conditions are met: (a) the knowledge or skill possessed is generalizable to the new situation and (b) the child perceives the applicability or utility of the knowledge or skill in the new situation.

In the present study, the correlational pattern between the English and Spanish reading measures suggest that a child's knowledge and skills associated with decoding are related across the two languages, as are those associated with overall reading ability, but to a lesser degree. This finding supports the premise that reading is a single process and that reading knowledge and skills gained in one language can be transferred, if the necessary conditions are met, to reading in another known language.

IMPLICATIONS: The practice of teaching children to read initially in their stronger language appears to be educationally sound. However, in commenting on the transfer of learning within a bilingual setting, researchers (e.g., Moll, Díaz, Estrada, & Lopes, 1981) contend that learning is primarily situation specific; generalizability to other situations depends upon whether the environment is organized to provide similar features that will facilitate its applicability to a different setting. Therefore, lesson environments, particularly as they relate to participant structures, have to be constructed in such a way that what children learn in Spanish reading class, for example, will be perceived as applicable in the English class and vice versa. Such being the case, this would suggest that there should be close coordination between those doing the instruction in the two languages. It further suggests that planning and teaching for transfer of learning should be included in the training of teachers who work with bilingual children.

INTEGRATION OF DATA SOURCES

The study assessed the degree to which various entry skills and instructional program indices could account for above or below average skill in each instructional year with respect to the set of component reading skills that were of primary interest (decoding, listening comprehension, and reading comprehension).

The predictor variables employed included: (a) entry language skill (based on teacher ratings), (b) task-specific entry skill (based on the previous year performance for the relevant task), (c) nominal instructional program (i.e., the number of years of Spanish reading prior to entry into exclusive English reading instruction), (d) indices of the quality and quantity of instruction received (based on factors derived from the classroom observation and teacher interview data), and (e) student attendance. In general, these variables were quite successful as predictors of skill, accounting for 75% to 95% of the variance in each of the nine reading measures in each of the four instructional years.

Based on the correlation and regression results, the predictor variables of kindergarten entry language skill, performance during the previous year, and the quantity and quality of reading instruction were most strongly related to reading achievement within each of the instructional years.

Summary of the English Relationships

- English kindergarten entry language skill is associated with above average performance in each of the IRAS component literacy skills assessed throughout the early grades.

- Knowledge of the English alphabet at kindergarten entry is strongly related to decoding-based literacy skills at first grade exit, independently of oral language entry.

- For a given literacy skill, entry skill is related to exit skill, and increasingly so over grade levels, thus suggesting that students tend to become academically "locked in place" with increased schooling.

- Enrollment in Spanish reading programs is generally negatively associated with acquired English literacy skill (but much of this relationship is due to entry level differences). There is some indication of relatively superior English literacy skills at fourth grade exit for those students with longer (longitudinal) enrollments in such Spanish reading programs (though the fourth-grade sample is limited).

- Relationships for both observed and planned instructional dimensions suggest that (a) literacy skills are advanced by instruction that makes strong formal language demands on students, by instruction that employs primary materials, and by instruction that engages students in work with text materials; and (b) comprehension skills and vocabulary skills are advanced by increased amounts of instructional time devoted to such skill development, but decoding skills show the opposite relationship, perhaps because of the relatively low quality of such instruction found in this data set.

- Literacy skills tend to show greater improvement with increased exposure to instruction -- the more opportunity for learning, the greater the skill acquired. This relationship was in many instances not generic in the present data set. Instead, more time on a particular component was correlated with growth in that component.

- Finally, some site contrasts are evident even after site differences due to entry skill have been removed, but these are relatively isolated.

Summary of Spanish Relationships

- As was true in the English data, Spanish language level on entry to kindergarten is associated with above average performance in each of the IRAS component literacy skills assessed during all of the early grades.

- Knowledge of the names of the letters in the Spanish alphabet is weakly (though positively) related to first grade exit performance, and, unlike the situation with the English alphabet, does not serve as a general index of pre-school literacy skill. However, in all subsequent instructional years, previous performance on a given IRAS scale is related to exit performance on that scale, increasingly so over grade levels. As in English, the relative standing of individual students with respect to literacy in Spanish becomes more rigid with increased schooling.

- Enrollment in Spanish reading programs is positively related to the acquisition of Spanish literacy in the early grades; this association becomes negligible in the later grade levels. Given the practice of transferring the most successful students (with respect to Spanish literacy skill) in Spanish reading programs, it is understandable that length in the program is not substantially related to acquired skill in these later grades.

- Relationships for both observed and planned instructional dimensions suggest that (a) literacy skills in general are advanced by instruction that engages students in work with text materials and by limiting interruptions and (b) decoding skills are advanced by increasing the quantity and quality of decoding instruction and decreasing the number of students in an instructional group.

- Attendance tends to be positively related to acquired literacy skill, but these relationships are weaker than those found in the English data.

- Site contrasts, even after entry skill differences have been removed, show that Spanish literacy skill is more advanced at those border sites that provide the greatest non-school support for Spanish.

ISSUE: Precursor Skills and Reading Achievement

As noted throughout this report, entry oral language and prereading skills are associated with reading achievement. When compared with children with less well developed skills at entry, children with relatively better developed oral language and prereading skills at entry were better able to take advantage of the instruction offered and to maintain their relatively superior level of attainment in reading throughout the primary grades.

IMPLICATIONS: Entry English language skills have pervasive and lasting effects on English reading achievement throughout the early elementary grades. While the oral English growth rate of the overall sample proceeded at a rate above the expectations of the growth track model, that of the Low English entry students showed a greater rate of growth than that of the High English entry group but did not converge with that of the High English entry group until late fourth grade. This suggests that the acquisition of "school-related" skills in a second language takes time. An important question for educators is, "What are effective intervention strategies for ensuring academic progress during the years while these children are in the process of gaining the necessary proficiency in English?"

An additional important challenge for educators is finding means to ensure, and perhaps accelerate, language and reading growth of students who at entry into school are deemed by their teachers to have relatively low level verbal skills. These are the children who got off to a slow start in school, gained somewhat less than a year of growth for a year of instruction, and fell further behind their more academically-prepared peers as they progressed through the early elementary grades.

ISSUE: Nominal Instructional Program and Reading Achievement

The study examined the degree to which the number of years students were enrolled in a Spanish reading program could account for reading achievement within each of the instructional years. For Spanish literacy, enrollment in Spanish reading programs is positively related to reading achievement during the early grades, but this relationship becomes negligible in the later grades. Children who are placed in these programs are generally those who are deemed by the schools to be limited in their English skills and to have stronger skills in Spanish than in English at the point formal reading instruction is begun, usually in first grade. These children remain in Spanish reading programs until they (a) reach a predetermined level of oral proficiency in English and (b) have attained a specified level of reading in Spanish and/or perform at or above a specified percentile score on a standardized test of reading achievement in English (usually the 40th percentile). In this study, some students in these programs received Spanish reading instruction for one year before being transferred to English reading; others remained in Spanish reading for two, three, or four years, with most being transferred to English reading by the end of third grade. Once transfer to English occurred, no further reading instruction in Spanish was provided, except during a brief "transition" period in some schools. With such criteria for transfer, the few students who remained in Spanish reading programs beyond the third grade were likely to have been children who were having trouble learning to read, since the oral English skills of most of the students by third grade exit tended to meet or exceed the oral English criterion for transfer.

While acquired English literacy skills were found to be generally negatively associated with numbers of years of enrollment in Spanish reading programs, there is some indication of relatively superior English literacy skills at fourth grade exit for those students with longer (longitudinal) enrollments in such Spanish reading programs. Although the sample was limited for this instructional year, this trend in the data raises some interesting questions. For children who begin initial reading instruction in Spanish, is there a threshold level that must be reached in Spanish reading for the benefits of such instruction to affect positively growth in English literacy? If so, does it correspond to the level of literacy that monolingual children normally achieve by the end of third grade? Are children in transitional bilingual education programs, where criteria for transfer to English reading is strongly tied to English performance, being kept in Spanish reading programs sufficiently long for them to attain the requisite literacy skills in Spanish? Does the time frame of this study, kindergarten through grade four, capture the long-term effects of initial reading instruction in the non-English home language? These and related questions merit attention as they are central to the current controversy surrounding transitional bilingual education.

ISSUE: Quantity and Quality of Instruction

To what extent does the quantity and quality of the instruction delivered to bilingual children affect reading achievement? Of the many factors that impact on student progress in reading, instruction is the one factor for which the schools have primary responsibility and which they have the most control. Therefore, identifying instructional patterns that are associated with success and failure, both in the early stages of reading instruction and in subsequent years, is a critical issue surrounding improvement of practices for all children.

Educational research over the last 10 to 15 years, conducted primarily with students from the general school population, has produced a well-founded knowledge base that allows educators to point with confidence to characteristics and actions that differentiate between instructional settings in which students successfully master the learning goals set out for them and those in which students are less successful. It identifies and describes what effective teachers do and how effective instruction is accomplished in effective schools. Similarly, some of the most eminent reading experts claim that the best teachers in the best schools know how to turn students into proficient readers (Anderson, Hiebert, Scott, & Wilkinson, 1985).

Research in bilingual education and related topics has also accumulated a substantial knowledge base in the last 10 years. The focus on academic achievement prior to and in the early part of the decade led to more and more investigations into the interaction between differences in the languages of instruction and the language of the student. This in turn uncovered a variety of variables that led to research into school and classroom climate, teacher and student

1996

variables, and pedagogical, socio-cultural, and legal issues. Thus, a considerable body of research exists that speaks directly to issues related to language minority education.

When examining the findings from the instructional data from the present study in relation to the knowledge base contained in the literature referred to above, one finds instructional patterns and teacher behaviors that are associated with (a) student academic gains in monolingual classrooms, (b) successful practices in bilingual classrooms, and (c) less reading gains in monolingual classrooms.

Practices Associated with Student Academic Gains

Factors present in the data that are associated with student academic gains and successful practices in both monolingual and bilingual classrooms include the following:

- o.. Strong focus on academic work; time spent working with textual materials (as opposed to time spent with non-textual materials).
- o.. Time allocated to reading and academic verbal interaction; literacy skills tend to show greater improvement with increased exposure to instruction -- the more opportunity for learning the greater the skill acquired.
- o.. Use of active teaching practices; relatively large amount of instruction from and interaction with the teacher.
- o.. High achievement expectations; use of tasks of appropriate difficulty level that challenge the students but allow consistent success.
- o.. Efficient classroom management; allocated instructional time devoted to instruction; classrooms that are relatively free of major behavioral disorders.

Additional factors associated with successful bilingual classrooms include the following:

- o.. Use of the home language with Limited English Proficient students some of the time.
- o.. Use of English primarily during English-medium instructional periods and Spanish primarily during Spanish-medium instructional periods.

Practices Associated with Less Student Gains

Factors associated with less student gains in reading in the present study and in research on monolingual students may be summarized as follows:

- Amount and quality of decoding instruction (inappropriate amounts or timing of such instruction; non-explicit instruction on letter-sound pairing).

- Limited attention given to explicit instruction to develop vocabulary (word meaning) and higher-order comprehension strategies (beyond those of comprehension of literal facts).

- Ability grouping of students, which may not be in the best interest of low achieving students. Children who get assigned to the lower groups get locked into an instructional track in which the range of instruction delivered is such that these students have limited opportunity to learn more than a narrow range of the skills and content needed to become fluent readers.

- Extensive use of seatwork assignments for low reading group students. Recent research suggests that seatwork is qualitatively a different experience for lower achievers than for high achievers. The two groups differ in terms of fluency of their answers and the appropriateness of strategies used. This may explain why achievement difference widens over time. Low achievers are spending less of their seatwork time in beneficial ways.

IMPLICATIONS: The classrooms in this study exhibited several of the characteristics of effective instruction, and for the students (in the aggregate) such instruction produced approximately a year of growth for a year of instruction in English reading comprehension as measured by performance based tests -- instruction makes a difference.

While similar in many ways, variation was noted among the classrooms on the quality of the dimensions of instruction assessed in the study. This suggests that to ensure effective instruction of all students, certain instructional dimensions need to be strengthened. Staff development should aim toward training teachers to (a) monitor their own use of language in the classroom and to provide instructional activities that make strong formal language demands on students; (b) make optimal use of textual materials, favoring these over non-textual materials in both direct instruction and independent work; (c) increase instruction in word meaning and the higher-order comprehension skills and to strengthen such skills through making this instruction explicit; and (d) evaluate the decoding needs of their students and to tailor their instruction on decoding to the identified needs, making such instruction explicit and limited to appropriate amounts. In addition, the practice of grouping students for instruction needs careful consideration, not only in terms of optimal size but also in terms of student membership, permanency of the group once formed, and instructional treatment provided.

ISSUE: Site Characteristics and Reading Achievement

The sites included in the study were selected to achieve variation on several dimensions (e.g., size, socioeconomic status, degree of urbanicity, concentration of Hispanic students, characteristics of the reading program). Given differing contextual environments, site differences in language and literacy development could be expected. For English, site contrasts in such development were relatively isolated, suggesting that schools were adjusting schooling practices and instruction to accommodate the needs of the local school population. Spanish literacy, on the other hand, was more advanced at certain of the border sites where substantial non-school support for Spanish was available.

IMPLICATIONS: Factors outside of the school play an important role in maintaining and/or fostering development of the non-English home language. Prominent among these are locale and the extent to which the language is used in the community and the wider environment, as well as the role of the home language in the affairs of the home and of the community; attitude of the student and others toward the maintenance of Spanish; and the extent to which written materials and formal usage are available to the students in the home language.

Without strong support from the home and the community, students in transitional bilingual education programs are not likely to achieve high levels of literacy in Spanish. Indications are, however, that these programs can, and are, promoting English literacy for all students. In this study, the students on the average were acquiring English oral language skills at the rate expected and were gaining in English literacy at or near a year of gain for a year of instruction, depending upon the type of reading measure used. Further, slightly more than half of the students were reading in English at grade level expectations by the end of second grade. Are these realistic expectations for schools to hold for students from non-English language backgrounds who at entry into school are limited in their English skills? Are growth rates of these youngsters similar to those generally reflected by mainstream school children under current schooling practices in public school systems in the United States? Do the growth rates of these youngsters compare favorably with those of their monolingual peers in the same school?

References

Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I.A. G. (1985). Becoming a nation of readers: The report of the Commission on Reading. Champaign, IL: University of Illinois - Becoming a Nation of Readers.

Calfee, R., & Associates. (1978). Stanford foundation skills test. Stanford, CA: Stanford University.

Calfee, R., & Associates. (1980). Stanford foundation skills test (rev. ed.). Stanford, CA: Stanford University.

Calfee, R., & Calfee, K. (1976). Reading and mathematics observation system - RAMOS/II (rev. ed.). Stanford, CA: Stanford University.

Calfee, R., & Calfee, K. (1978). Reading and mathematics observation system - RAMOS/SEDL. Austin, TX: Southwest Educational Development Laboratory.

Calfee, R., & Calfee, K. (1979). Interactive reading assessment system. Stanford, CA: Stanford University.

Calfee, R., & Calfee, K. (1981). Interactive reading assessment system (rev. ed.). Stanford, CA: Stanford University.

Calfee, R., Calfee, K., & Peña, S. (1979). Interactive reading assessment system - Spanish. Stanford, CA: Stanford University.

Calfee, R., & Peña, S. (1978). Prueba Stanford de destrezas fundamentales. Stanford, CA: Stanford University.

Calfee, R., & Peña, S. (1980). Prueba Stanford de destrezas fundamentales (rev. ed.). Stanford, CA: Stanford University.

Consulting Psychologists Press, Inc. (1971). Group embedded figures test. Palo Alto, CA: Author.

Cummins, J. (1983). Policy report: Language and literacy learning in bilingual instruction. Report to the National Institute of Education. Austin, TX: Southwest Educational Development Laboratory.

DeAvila, E. (1976). Cartoon conservation scales. Corte Madera, CA: Linguametrics Group, Inc.

DeAvila, E., & Duncan, S. (1977). Language assessment scales. Corte Madera, CA: Linguametrics Group, Inc.

Doebler, L. K., & Mardis, L. J. (Winter 1980-81). Effects of a bilingual education program for Native-American children. NABE Journal, 5(2), 23-28.

Duncan, S., & DeAvila, E. (1976). Student operational language assessment scale. Austin, TX: Southwest Educational Development Laboratory.

González, G. A. (1977). Brownsville Independent School District bilingual education program Title VII - Final report for 1976-77. Brownsville, TX: Brownsville Independent School District.

Hispanic Policy Development Project. (1984). Make something happen: Hispanics and urban school reform. Vol. 1, 2. Washington, DC: Author.

Kagan, J., Rosman, B., Day, D., Albert, J., & Phillips, W. (1964). Information processing in the child. Psychological Monographs, 78, 1-37.

Klee, C. (1984). A discourse analysis of the oral language interactions of Spanish/English bilingual children in three environments. Unpublished doctoral dissertation, University of Texas, Austin.

Leyba, C. F. (1978). Longitudinal study of a Title VII bilingual program, Santa Fe Public Schools, Santa Fe, New Mexico. Los Angeles, CA: National Dissemination and Assessment Center.

Mace-Matluck, B., Tunmer, W., & Domínguez, D. (1979). Oral language proficiency rating scale. Austin, TX: Southwest Educational Development Laboratory.

Moll, L. C., Díaz, E., Estrada, E., & Lopes, L. M. (1981). The construction of learning environments in two languages. Monograph. San Diego, CA: Laboratory of Comparative Human Cognition.

O'Malley, J. M. (1982). Children's English language and services study: Educational needs assessment for language minority children with limited English proficiency. Rosslyn, VA: National Clearinghouse for Bilingual Education.

Rosier, P., & Farella, M. (1976). Bilingual education at Rock Point - some early results. TESOL Quarterly, 10, 379-388.

San Diego City Schools. (1982). ESEA Title VII bilingual demonstration project. San Diego, CA: Author.

Southwest Educational Development Laboratory. (1978). Reading Teacher checklists. Austin, TX: Author.

Southwest Educational Development Laboratory. (1979). Bilingual classroom questionnaire. Austin, TX: Author.

1101

Southwest Educational Development Laboratory. (1980). Survey of teachers' background and language skills. Austin, TX: Author.

Southwest Educational Development Laboratory. (1981). Inventory of bilingual instruction. Austin, TX: Author.

Troike, R. C. (1981). Synthesis of research on bilingual education. Educational Leadership, March, 498-504.

Willig, A. C. (1985). A meta-analysis of selected stuaies on the effectiveness of bilingual education. Review of Educational Research, 55(3), 269-317.

Witkin, H., Oltman, P., Raskin, E., & Karp, S. (1971). A manual for the embedded figures tests. Palo Alto, CA: Consulting Psychologists Press, Inc.

1102