MCCALL-PEREZ, FREDERICK CLANCY
THE PERFORMANCE OF BILINGUAL CHILDREN ON THE SPANISH STANDARDIZED ILLINOIS TEST OF PSYCHOLINGUISTIC ABILITIES.

THE UNIVERSITY OF ARIZONA, PH.D., 1979

THE PERFORMANCE OF BILINGUAL CHILDREN ON THE SPANISH STANDARDIZED ILLINOIS TEST OF PSYCHOLINGUISTIC ABILITIES

by

Frederick Clancy McCall-Perez

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF SPECIAL EDUCATION

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 7 9

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

I hereby recommend that this dissertation prepared under my direction by Frederick Clancy McCall-Perez

entitled The Performance of Bilingual Children on the Spanish Standardized Illinois Test of Psycholinguistic Abilities

be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Dissertation Director — Date 3/27/79

As members of the Final Examination Committee, we certify that we have read this dissertation and agree that it may be presented for final defense.

Samuel A. Kirk — Date 3/27/79

Henry E. Butler Jr. — Date March 12, 1979

Gregory F. Olma — Date 3-12-79

Date

Ray A. Blake — Date 3/13/79

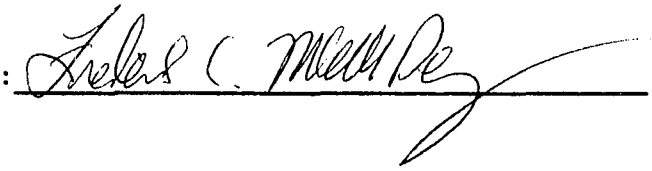
Final approval and acceptance of this dissertation is contingent on the candidate's adequate performance and defense thereof at the final oral examination.

11/78

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: 

ACKNOWLEDGMENTS

The author would like to thank all the members of his doctoral committee: Dr. Gregory Aloia, Dr. Roy Blake, Dr. Henry Butler, Dr. Samuel A. Kirk, Dr. Jeanne McCarthy.

I would especially like to thank Dr. Jeanne McCarthy for her patience in guiding my development. Dr. Henry Butler is someone whose openness to new ideas and understanding I admire and thank him for. I admire the wisdom of Dr. Samuel Kirk and appreciate his relentless encouragement to be a professional.

The author thanks Hayward Unified School District in Hayward, California for their support of this study. I especially thank Dr. Bushnell and Mr. Lawrence for their interest in the research and help in making the bilingual programs and population available to me.

Finally, I thank my friends and family for their emotional support when I most needed it.

TABLE OF CONTENTS

TABLE OF CONTENTS--Continued

LIST OF TABLES

LIST OF TABLES--Continued

ABSTRACT

The purpose of this investigation was to determine if bilingual children score significantly higher on a testing instrument normed in Spanish or an English translation of the same test. Twenty-six kindergarten and 28 second grade children enrolled in school in a California community were tested on the Illinois Test of Psycholinguistic Abilities standardized on a monolingual Hispanic population and its English translation. All children were Spanish-American and reported that Spanish was spoken in their homes; all were in bilingual programs identified by teacher judgment as "bilingual." Twenty-eight percent of the subjects were from middle socioeconomic status families and 72 percent were from lower socioeconomic status families.

In testing these subjects on the Spanish standardized ITPA and its translation, the split-half method was used. Half of the children at each grade level were tested first in Spanish and then tested within four days in English. The other half of the subjects were tested in English first and the Spanish test was administered within four days.

The results of this study were:

1. Bilingual children scored as well or better on the English translation of the Spanish ITPA. The kindergarten children scored significantly higher on the English translation on two of the ten subtests. The second grade children scored

significantly higher on the English translation on four of the ten subtests and on the total score.

2. The differences between the performance of the children on the English and Spanish tests appeared to be primarily on tests in the auditory-vocal channel rather than on tests in the visual motor channel.
3. Mean deviation scores were calculated for all subjects as a possible criterion of a specific learning disability. At the kindergarten level this criterion identified two children on the English translation and none on the Spanish version. At the second grade level, however, this criterion selected six children when the Spanish version was used, but only one when the English translation was used. In other words, by second grade there were more children with wide discrepancies of scores on the Spanish version.
4. The Hayward sample tested lower than the monolingual standardization sample, but similar to the bilingual Tucson-San Diego sample.

It should be emphasized that the results of this study apply only to the sample of bilingual Hispanic children from a bay area city, and they should not be generalized to Hispanic bilingual children whose Spanish is dominant. They should not be generalized to bilingual children who are significantly more fluent in Spanish.

CHAPTER 1

INTRODUCTION

The purpose of this investigation is to determine if bilingual children score significantly higher on a testing instrument normed in Spanish or an English translation of the same test. Current interest in the problem of testing bilingual children developed because of recent legislative and court action. There has been increased pressure on the public school systems to test non-English speaking and bilingual children in their native language and to make educational decisions based on testing done in the child's primary language. In California a recent court case (Diana v. State Board of Education, 1970) required the retesting of over 22,000 Spanish-speaking children. The courts banned the use of English intelligence tests for Spanish-speaking children saying they unfairly classified some children as handicapped who were not.

Three types of legal action have contributed to the requirements that schools test children in their primary language: (1) federal legislation, (2) state legislation, and (3) court cases.

Federal Legislation

Legislative policy concerning the education of bilingual children took a new direction in 1965 with the enactment of the Bilingual Education Act of 1965.

> In recognition of the special educational needs of the large number of children of limited English-speaking ability in the United States, Congress hereby declares it to be the policy of the United states to provide financial assistance to local educational agencies to develop and carry out new and imaginative elementary and secondary school programs designed to meet these special needs. For the purpose of this title, "children of limited English-speaking ability" means children who come from environments where the dominant language is other than English (Bilingual Education Act, Title VII, 1965, p. 49).

Although the Bilingual Education Act of 1965 makes no direct statements about testing of bilingual children, this act and subsequent amendments (Bilingual Education Act of 1968, 1970) authorize distribution of federal moneys to local school districts to facilitate implementation of diverse programs. For many of these programs it would be necessary to test the level of skills or potential of the child.

A memorandum dated July 17, 1970 from J. Stanley Pottinger (Swanson, 1974), Director, Office of Civil Rights, further exemplifies urgency by the federal government to have school districts respond to bilingual children's needs in a special way.

The purpose of the memorandum was to clarify HEW's policy on issues concerning the responsibility of school districts to provide equal educational opportunity to national origin-minority group children deficient in English language skills. The memorandum listed four major areas of concern, which would be regarded as part of the compliance responsibilities under Title VI of the Civil Rights Act of 1964

> (1) Where inability to speak and understand the English language excludes national origin-minority group children from effective participation in the educational programs offered by a school

> district, the district must take affirmative steps to rectify the language deficiency in order to open its instructional programs to these students.
>
> (2) School districts must not assign national origin-minority group students to classes for the mentally retarded on the basis of criteria which essentially measure or evaluate English language skills; nor may school districts deny national origin-minority group children access to college preparatory courses on a basis directly related to the failure of the school system to inculcate English language skills.
>
> (3) Any ability grouping or tracking systems employed by the school system to deal with the special language skill needs of national origin-minority group children must be designed to meet such language skill needs as soon as possible and must not operate as an educational dead-end or permanent track.
>
> (4) School districts have the responsibility to adequately notify national origin-minority group parents of school activities which are called to the attention of other parents. Such notice in order to be adequate may have to be provided in a language other than English (Federal Register, 1970, p. 11595).

The memorandum in conjunction with the Civil Rights Act of 1964 has the force and effect of law and applies to all school districts having five percent or more national origin-minority group children (Civil Rights Digest, 1971). School districts who fail to meet the memorandum's compliance requirements may have all of their federal funds withdrawn.

Importantly, in this memorandum, testing of bilingual children is clearly referred to in paragraph two. The message is simple: for special class placement bilingual children must not be tested on instruments that evaluate English language skills. The bilingual child must be evaluated in his primary language before placement decisions are made.

Stimulated by federal action, state authorities, too, have increased requirements for bilingual programs which include the assessment of bilingual children in their primary language.

State Legislation

Prior to the enactment of the Federal Bilingual Education Act of 1968, state legislation was not concerned with most bilingual education. Many states actually prohibited bilingual education and required all instruction and testing to be done in English. Seven states had imposed criminal penalties or revocation of a teacher's certificate for teachers who gave non-English instruction (Kobrick, 1972).

The Arizona Revised Statutes of 1973 extended "permissable" coverage through 1981. The new law stated "where there are pupils who have difficulty in writing, speaking or understanding the English language because they come from an environment wherein another language is spoken primarily . . . the district may provide special programs" (Arizona Revised Statutes, 1973, p. 1098D).

In general, state legislation is becoming friendly towards bilingual education and assessment. Most states encourage school districts to initiate bilingual programs at their discretion and encourage testing children in their primary language. In many states it is standard practice to test children in their primary language.

At present, two states, Massachusetts and Alaska, require certain school districts to institute bilingual programs which include assessment in the child's primary language. The Alaska law goes so far as to state that the absence of bilingual programs has created a

great learning handicap for those to whom English is a second language.

It is clear that federal and state action has increased pressure to start new programs for the bilingual child and to test bilingual children in their primary language. Additional pressure on school districts to test children in their primary language has come from recent court decisions.

Court Decisions

Several court decisions in recent years have required school districts to create special programs for bilingual children including testing in the child's primary language. The proponents of bilingual education in these cases have alleged that the school's failure to provide such programs is a denial of equal educational opportunity in violation of the Equal Protection Clause of the 14th Amendment, claiming discrimination based on national origin, a violation of a fundamental right to education. The opponents have countered that the schools are required to provide the same education to all. What follows is a brief review of the important cases which promote bilingual programs and the testing of children in their primary language.

Lau et al. v. Nichols et al. (1973)

A landmark case that influenced recent court decisions nationally is Lau et al. v. Nichols et al. (1973). This case was a class action suit on behalf of all non-English speaking Chinese students in the San Francisco Unified School District. It is by far the best known case dealing

with bilingual children and has had nationwide implications. The suit contended that the school district had abridged bilingual children's rights to an education and to bilingual education, and disregarded their rights to equal educational opportunity among themselves and with English-speaking students. The plaintiffs asked for mandatory bilingual education for all primary non-English-speaking Chinese students. Originally, the district court denied all relief to appellants, claiming appellees had no duty to rectify appellants' special deficiencies, as long as they provided these students with access to the same educational system made available to all other students. In a 2-1 decision, the Ninth Circuit Court of Appeals affirmed the district court. The plaintiffs argued, however, that Brown v. Board of Education (1954) mandated consideration of a student's responses to the teaching provided by his school in determining whether he had been afforded equal educational opportunity. It was argued also that Brown v. Board of Education required schools to provide "equal" opportunities to all, and equality was to be measured not only by what the school offers to the child, but by the potential that the child brings to the school. If the student was disadvantaged with respect to his classmates, the school, it was contended, had an affirmative duty to provide him special assistance to overcome his disabilities, whatever the origin of those disabilities might be. These arguments were rejected.

In 1974 on appeal the Supreme Court unanimously reversed the circuit court decision in its amicus brief on the Lau case holding that the failure of the San Francisco school system to provide English language instruction to approximately 1,800 students of Chinese

ancestry who do not speak English denies them a meaningful opportunity to participate in the public educational program and thus violates 601 of the Civil Rights Act of 1965.

Although the court avoided deciding the case under the Equal Protection Clause of the 14th Amendment, it is to the Justice Department's credit that in its amicus brief in Lau v. Nichols that it forcefully argued the alternate approach under the Civil Rights Act to the court. In this decision for the first time the Supreme Court ruled favorably for bilingual education and committed itself to assist if enforcement difficulties developed.

The court recognized that failure to provide some form of compensatory English instruction to bilinguals deprived them of a meaningful and effective educational opportunity, dispelling the notion that identical treatment will suffice in providing educational opportunity, at least where the students are not equal in the competencies they bring to school. The court also recognized the denial and the responsibility of the school to take affirmative action even though it did not cause the student's deficiency. The Lau v. Nichols decision was the most important court decision in reference to bilingual children and has had far-reaching effects on bilingual education. This decision influences many court decisions, some of which will be mentioned here.

Guadalupe Organization Inc., v.
Tempe Elementary School District (1972)

In this important decision, the Tempe elementary special education program was required to retest bilingual mentally handicapped

children to determine if any bilingual children had been incorrectly assigned to special education. The court ruled that children had to be retested in their primary language. Further, the records of any children incorrectly placed in special education were ordered by the court to be destroyed. The court also required that any communications to parents of these children be in the family's primary language. This court action was prompted by a disproportionate number of bilingual children being enrolled in special education classes. This decision is the best known to date, is quoted in many decisions since 1972, and has provided a great deal of pressure on school districts across the country to rely on non-English assessment instruments.

Serna v. Portales Municipal Schools (1973)

In Serna v. Portales Municipal Schools, a class action by the parents of Spanish-surnamed minority students, it was claimed that discrimination was a result of an educational program within the school district that was tailored to educate the middle-class child from an English-speaking family without regard to the educational needs of the child from an environment where Spanish was the primary language spoken. The district court found that Spanish-surnamed minority students consistently scored lower on achievement tests. The school district argued that the special educational needs of Spanish-surnamed minority students were not the result of state action and that it did not create the language problems through any classification or racially motivated discrimination. The school district had relied on Keys v. School District (1973) for its contention. The district court read the Keys opinion to

indicate that, while the court was unwilling to hold that segregation of itself would be a denial of equal educational opportunity, it would be a deprivation of equal protection for a school district to effectuate a curriculum that was not tailored to the educational needs of minority students. The district court held that the Portales school district's curriculum failed to take into account the educational needs of the Spanish-surnamed minority students, constituting state action and depriving such students of an equal educational opportunity in violation of the Equal Protection Clause of the 14th Amendment. The district court ordered the school district to implement adequate bilingual education programs in all of its elementary schools. This influential decision put direct pressure on this school district to provide special programs for bilingual children. It also put indirect pressure on this district and others to develop means to assess children in their primary language.

In Diana v. State Board of Education (1970), the court banned use of an English intelligence test for Spanish-speaking children. This class action suit alleged that such test items as "What is a chattel? What does C.O.D. mean?" unfairly assessed Spanish-speaking students (many of whom are bilingual) and resulted in their being placed in special education classes in disproportionate numbers. The court saw this as "tantamount to a life sentence of illiteracy and public dependency" (Diana v. State Board of Education, 1970, p. 32).

The complaint alleged that bilingual children comprised 26% of California's students in classes for emotionally handicapped and

mentally retarded while they comprised only 13% of the school-age population. This court decision had immediate ramifications for all of California and served as an impetus for school districts throughout the nation to test bilingual children on tests of intelligence in their native language.

United States v. Texas (1971)

The district court in the above case found that Del Rio, Texas maintained a dual educational system, keeping separate and apart the Mexican-American student from the Anglo student. The court ordered the integration of the two school systems, creating a unitary school system with no Mexican schools and no Anglo schools, but a system of integrated schools. The court found that Mexican-Americans constituted an identifiable ethnic or national origin minority group in the State of Texas, and that Mexican-Americans are subject to protection under Title VI of the Civil Rights Act of 1964 and the 14th Amendment as applied to racial and ethnic discrimination in public schools. The finding of state action was based on the financing of the segregated systems with state funds without which neither system could have sustained itself. The court stated that mere integration was not sufficient and that special educational consideration must be given to the Mexican-American students in assisting them to adjust to those parts of their new school environment which presented a cultural and linguistic shock. The court included in its ordered integration a comprehensive educational plan for the consolidated school district, including a broad bilingual-bicultural program, recruitment of bilingual faculty,

curriculum design, funding and community participation. The aim of the plan was to decrease the high school dropout rate of minority students by providing specialized instruction compatible with cultural and learning characteristics. The court concluded by stating that there was a

> need to avoid the creation of a stigma of inferiority akin to the "badges and indicia of slavery." To avoid this result the Anglo-American students too must be called upon to adjust to their Mexican-American classmates, and to learn to understand and appreciate their different linguistic and cultural attributes. The process by which all students participate in a joint learning and adjustment process will not only constitute an educational enrichment but also, will bring the school system as a whole closer to that goal or state-of-being referred to by the Supreme Court as a "unitary system" (p. 649).

Aspira of New York, Inc. v.
New York Board of Education (1972)

Aspira of New York, Inc. v. New York Board of Education was a class action suit on behalf of 170,000 New York City public school students (primarily Puerto Rican) whose primary language was not English and who receive no services or inadequate services to meet their linguistic needs. The suit asked that bilingual educational programs be required for bilingual children. The plaintiffs argued that the present educational program resulted in high levels of illiteracy, truancy, and unemployment among bilingual children, and had contributed to their low scores on standardized reading tests and underrepresentation in high school graduation classes. Plaintiffs claimed that this was in violation of the 1st, 5th, 9th, and 14th amendments and Title VI of the Civil Rights Act of 1964. The court ordered New York to start bilingual programs for all bilingual students in the district.

In May of 1975 because of unnecessary delay in carrying out this order, the court further ordered that all Spanish-surnamed children receive a Spanish test battery rather than those children just falling within a certain percentile as requested by the district. Children may be tested in both languages.

Summary

This chapter introduced the purpose of the study, to determine if bilingual children score differently on a testing instrument normed in Spanish then they do on its English translation. It is obvious that legal action has pressured the schools to test children in their primary language. Unfortunately, court decisions have been based on opinion rather than on scientific knowledge. Courts have not been presented with scientific evidence that bilingual Hispanic children score higher on tests administered in Spanish. Instead, legal action reflects societal opinion--the "trend of the times"--without first looking at empirical evidence. The empirical evidence which exists regarding the performance of bilingual children on intelligence tests will be reviewed in Chapter 2.

CHAPTER 2

REVIEW OF LITERATURE

This review will cover studies of the performance of bilingual children on tests given in the primary language or in English. Before reviewing studies, it is important to acknowledge some of the problems in doing research with bilingual subjects. Three problems will briefly be mentioned here: (1) difficulties in determining the degree of bilingualism, (2) lack of availability of Spanish-standardized instruments, and (3) variability in the types of tests used.

Determining the Degree of Bilingualism

Many of the early researchers completely ignored the problem of defining bilingualism or of using an objective measure to determine bilingualism. Leopold (1939) said bilingualism is present when two languages are used as a media of discourse. The court in Aspira v. New York Board of Education (1975) implied that all Spanish-surnamed children are possibly bilingual and required all Spanish-surnamed children to be tested in Spanish. Barbe (1933) defined a bilingual child as one who uses the mother tongue while young but acquires a second language from the environment. In reviewing existing tests to determine the degree of bilingualism, Ortiz (1979) points out that at present no reliable and valid instrument exists.

Lack of Spanish-standardized Instruments

The greatest hindrance in research dealing with bilingual subjects is that most researchers used translations of instruments standardized in English. The subjects being tested were dissimilar to the population the instrument was standardized on, making any interpretations questionable. Frequently the translated version differed in level of difficulty of the items.

Types of Tests Used

Various types of instruments have been used to measure intelligence and achievement of bilingual subjects. Results have varied depending on the type of instrument used. Questions need to be asked about the performance of bilingual children on various types of tests used. This review will first address the performance of bilingual children on group tests of intelligence. The studies utilizing individual verbal and nonverbal instruments will be examined. Finally, studies of the performance of bilingual children on tests of achievement will be reviewed. Since the majority of studies done were early studies (prior to 1950) and used group intelligence tests, that question will be addressed first utilizing representative studies.

The Performance of Bilingual Children on Group Tests of Intelligence

Colvin and Allen (1923) selected 50 children of American parentage and 50 children of Italian parentage in grades 5 through 8. No other measure of bilingualism was used. Colvin and Allen selected children that were of equal intelligence according to the school

records. All children were tested on a group intelligence test, The National Intelligence Test. The average for the American group was 103; the Italian group average was 90. Colvin and Allen felt that the Italian group scored lower on the National tests because they suffered from a language handicap.

In a similar study, Mead (1927) administered the Otis Group Intelligence Test to 276 children of Italian parentage and 160 children of American parentage. Mead implied that the Italian parentage children spoke Italian at home although no measure was mentioned. All children were in grades 6 through 10 in New Jersey public schools. The American children's group mean was 27.33 points higher than the bilingual children. It should be remembered that all children were tested in English. Mead felt that intelligence test scores achieved by foreign-parentage children were seriously affected because of a language handicap.

Pintner (1923) sought to determine whether a verbal group test would be a valid measure of intelligence of foreign children. He administered the National Intelligence Test and the Pintner Non-Language Test to all the children in the third and fourth grades of a New York City school. The children were divided into foreign and American groups. Membership in the foreign or American group was decided solely on the basis of parents' nationalities. The groups were not matched on any measure except grade in school. For all foreign children combined, the median MA on the Pintner Non-Language Test was the same as the median MA of the Americans on that test. On the National Intelligence Test, only 37 percent of the foreign group reached the median MA of the

American group. Pintner cautioned that foreign children may be penalized when verbal tests of intelligence were used as the sole criterion of measurement.

Hill (1936) conducted a study to determine the effect of bilingualism upon the measured intelligence of children of Italian parentage. Bilingual children of the first, third, and sixth grades in a New Jersey elementary school were matched with monolingual children with respect to chronological age, sex, mental age, IQ, and socioeconomic status. A subject was considered bilingual if he heard and spoke Italian at home.

On the first grade level, 36 monolinguals were compared with 36 bilinguals. On the third grade level, 36 monolinguals were matched with 36 bilinguals. The Haggerty Intelligence Test was used to match the groups as to mental age. The matched groups were then given the Pintner Non-Language Test, List 8 of the Morrison-McCall Spelling Scale, and the Wittmer Cylinder Test. On the sixth grade level, 50 bilingual children who were paired with 50 monolingual children were given the Otis Self-Administering Examination, Form A, List 8 of the Morrison-McCall Spelling Scale, the Army Beta Group Examination, and the Wittmer Cylinder Test.

Hill found no reliable differences in scores on verbal, nonverbal, and performance tests between the Italian children who heard and spoke Italian at home and Italian children who heard and spoke English at home at all three levels. He concluded that bilingual children are not penalized by use of English standardized tests. It must be remembered that Hill first matched his groups on intelligence tests administered in English before he began his investigation.

Pintner (1932) conducted an investigation in which the Pintner-Cunningham Primary Mental Test and the Pintner Primary Non-Language Test were administered to 430 children in three lower elementary grade schools in New York City. Children's surnames and teacher referral determined if a subject was placed in the bilingual or monolingual group. In one school (bilingual Italian), the mean score of the bilingual group on the Pintner-Cunningham Test was 8.3 points less than the mean score of the monolingual group. On the Pintner Non-Language Test, however, the bilingual group had a mean score only 4.02 points lower than the mean score of the monolingual group. In school two (bilingual Bohemian), the bilingual group exceeded the monolingual group on both tests. In school three no difference was found but Pintner felt this was probably due to the fact that no real division existed in the language background of the pupils in this school. It should be noted that none of the groups were matched as to socioeconomic status and there was a very general and subjective determination of bilingualism. The author concluded that with tests verbal in content, a greater handicap for non-English children is probable. The language handicap will presumably be greatest early, diminishing as the child grows older.

Hirsch (1926) conducted an extensive investigation of subjects attending the public schools in Massachusetts. This review of the investigation will be limited to that phase of the study which treats the differences in the intelligence test results found to exist between the 1,030 American and the 350 Italian subjects in grades one through nine. The "bilingual" (Italian) subjects were determined by parents' nationality.

The Pintner-Cunningham Primary Mental Test was administered to the subjects in the first grade; the Dearborn Group Test of Intelligence was administered to the subjects in the second through ninth grades. The mean IQ of the Italian children who took these tests was 85.8. The mean IQ of the American-whites was 98.3. The difference in mean IQ between the American-white and the Italians was 12.5 in favor of the Americans and was found to be statistically significant.

Hirsch concluded that bilingual children are penalized on the tests but cautiously does not attribute this to a language handicap factor. The fact that Hirsch's nationality groups were not controlled with respect to number, sex, age, or degree of bilingualism, and only to a minor extent as to socioeconomic status, must be taken into account when his results are being studied.

Rigg (1928) investigated differences on intelligence and achievement tests between a foreign group and American subjects in early elementary school in St. Louis. Subjects were in the foreign group if they answered a question, "What language is spoken in your home?" with anything but English. Groups were not matched on socioeconomic status. The National Intelligence Test, the Woody-McCall Arithmetic Fundamental Test, and the Thorndike-McCall Reading Tests were administered to all subjects. Rigg found the median IQ of the American subjects to be 104.85 and 91.43 for the foreign group. Rigg concluded bilinguals are penalized on English tests.

Jamieson and Sandiford (1928) studied approximately 700 bilingual Indian subjects in Southern Ontario, comparing them to monolingual children. The National Intelligence Test, the Pintner-Cunningham

Primary Mental Test, the Pintner Non-Language Scale, and the Pintner-Patterson Performance Scale were administered. For all of these tests except the Pintner-Patterson Performance Test, the monolingual group did significantly better than the bilingual group. This study was not controlled for socioeconomic status or degree of bilingualism. Jamieson and Sandiford stated that these results pointed to a severe language handicap for bilinguals since the monolingual subjects surpassed the bilinguals on all verbal tests.

Barbe (1933) investigated 395 bilingual subjects ages 10 to 14 selected from three mixed bilingual schools where Welsh was the dominant language. A subject was considered bilingual if he spoke Welsh at home before coming to school. A monolingual group was composed of 302 subjects from 10 to 14 years of age, who were chosen from two schools of equal socioeconomic status.

All subjects in the monolingual and bilingual groups were given the Pintner Non-Language Intelligence Test and the Northumberland Standardized Test. The Northumberland Test, which is a group verbal test of intelligence, was used chiefly as a check upon the nonverbal test. The results indicated that the monolingual group had an average superiority of .8 of a year on the Northumberland Test, whereas on the Pintner Non-Language Test, the bilingual group had an average superiority of .44 of a year. Barbe concluded bilingual children are penalized on verbal intelligence measures and advocated the use of a nonverbal test of intelligence when testing bilingual subjects.

Arsenian (1945) examined the relationship between bilinguals and the measurement of mental development of children between the ages

of 9 and 14 years in New York. The two main experimental groups consisted of 1,152 American-born Italian and 1,196 American-born Jewish children. Bilingualism was measured by means of the Hoffman Bilingual Schedule and seven questions from the Simms Score Card were used to measure socioeconomic status. The relationships between intelligence as measured by the Pintner Non-Language Test and the Spearman Visual Perception Test, Part 1, and age, sex, race, and socioeconomic status were also determined. The children of the experimental groups were divided into low and high bilingual groups on the basis of the Hoffman Bilingual Schedule at each age from 9 to 14. This was done to determine if there were differences in the mental development of low and high bilingual children. Arsenian matched a group of 38 monolinguals with a group of 38 bilinguals person for person, as to race, sex, socioeconomic status, and age in months to determine if there existed any reliable differences in intelligence between them. The Pintner Non-Language Test and the Spearman Visual Perception Test, Part 1, were used to measure intelligence of these groups. Arsenian found no relationship between bilingualism and intelligence for the groups studied. No reliable difference was found between the low and high bilingual groups as far as the measurement of mental ability was concerned. Arsenian concluded that bilingualism did not influence favorably or unfavorably the mental development of the children in his investigation.

Yoshioka (1929) administered the National Intelligence Test to 38 bilingual Japanese children in California schools. The subjects were in grades 3 through 8. A child was considered bilingual if the primary language of the home was Japanese. All subjects were

administered the English version of the test and a Japanese translation. Group A consisted of 17 subjects in grades 3 through 5 and was administered the Japanese translation on day one and the English test on day two. Group B consisted of 21 subjects in grades 6 through 8 who received the tests in reverse order. Yoshioka found that the subjects scored below norms for their age on both the Japanese and English test. He concluded that bilingualism is a hardship and the bilingual child is penalized on general intelligence tests.

Mitchell (1937) selected 236 Spanish-speaking pupils in early elementary grades. Teachers' judgment was used as the criterion for Spanish speaking. The Otis Group Intelligence Test was administered in English and in Spanish to all the subjects. The mean IQ at each grade was higher for the Spanish version than for the English version. For the three grades combined, the mean IQ was 9.28 points higher when the test was administered in Spanish. Mitchell found bilingual children to be penalized on English intelligence tests and concluded that bilingual children work under a serious handicap especially in the lower grades of elementary school. Mitchell felt the difficulty appeared as a general language handicap rather than in any specific phase an intelligence test might measure.

Stark (1940) tested monolingual and bilingual Irish subjects with the Dawson Mental Test. It is unclear what criteria Stark used to determine bilingualism. Elementary school age monolingual children were tested with English tests while bilingual subjects received the Irish version. Stark found mixed results at different age groups. It must be remembered the bilingual subjects did not receive the English

test. The results of this study are questionable because of small samples, lack of control for socioeconomic status, and no objective measure of bilingualism.

Discussion

The majority of the above studies indicate that bilingual children are penalized on group administered intelligence tests. Based on the results of these early studies, one could conclude that intelligence testing using group tests does penalize the bilingual child. However, the research thus far is not too helpful in determining if bilingual children perform better on tests given in English or the child's primary language. These early studies with the exception of Yoshioka (1929) and Mitchell (1937) tested bilingual children only in one language. This, combined with some of the individual criticisms of the studies, adds to the difficulty of interpreting the results. Second, and most important, most researchers would agree that group intelligence tests penalize not only the bilingual child but any child. Group intelligence tests are not as statistically sound as individually administered intelligence tests.

Conclusion: Bilingual children are penalized on group administered intelligence tests. It now becomes important to look at research dealing with bilingual children utilizing individually administered intelligence tests. Since there are two types of individual tests, verbal and nonverbal, studies utilizing individual verbal and nonverbal instruments will be reviewed.

The Performance of Bilingual Children on Individual Verbal and Nonverbal Intelligence Tests

Much of the research that exists concerning the performance of bilingual children on individually administered intelligence tests tested children in only English and many of these studies were not concerned with Spanish/English bilingualism. A representative sample of these studies will be reviewed first.

Seidl (1937) selected 240 subjects, all of whom were American born of Italian descent, from public schools in New York City. He included 120 monolingual children and 120 bilingual children, ages 9 through 11. A language questionnaire was used to determine the extent of the subjects' use of Italian. If it was high, the subject was placed in the bilingual group. Subjects were not matched person to person. Seidl tried to determine the effect of bilingualism on the estimates of intelligence obtained from the 1916 revision of the Stanford-Binet Scale for the verbal test, and the Arthur Point Scale of Performance Tests as the nonverbal scale. All subjects were tested in English only. Results showed that the monolinguals were superior to the bilinguals on all verbal tests. The average IQ of the monolinguals on the Stanford-Binet Scale was 96.25 compared to the average of 91.61 for the bilinguals. On the performance tests, however, the bilinguals were superior to the monolinguals. The average IQ of the bilinguals on the Arthur Scale was 100.41, as against the average IQ of the monolinguals of 96.21. Seidl concluded that the difference between the two groups on the verbal test could be assigned to the language handicaps of the bilingual subjects.

Darcy (1946) did an early study of the effects of bilingualism on the measurement of intelligence of preschool children of Italian parentage. A child was considered bilingual if the child heard and spoke Italian at home always or most of the time and heard and spoke English outside the home always or most of the time. The Stanford-Binet Scale was used and the Atkins Object-Fitting Test was used as a nonverbal means of assessment. All children were tested in English only. Darcy found a statistically significant difference between the mean IQ's achieved by the monolingual and bilingual subjects on the Stanford-Binet. Monolingual children scored significantly higher than bilinguals. The bilingual group, however, did significantly better than the monolingual group on the Atkins Object-Fitting Test. Darcy concluded that the bilingual subjects suffered from a language handicap.

Levinson (1959) compared the performance of bilingual and monolingual Jewish preschool children ages 5 to 6 on four intelligence tests to determine the test or test items which would be most suitable for appraising the intelligence of the bilingual subjects. Bilingualism was determined by parents' judgment. The Stanford-Binet Scale, the Wechsler Intelligence Scale for Children, the Progressive Coloured Matrices, and the Goodenough Draw-A-Man Test were administered to 57 monolingual and 60 bilingual subjects in two sittings in counterbalanced order. Groups were matched by fathers' occupation.

Significant differences at the five percent level of confidence were found between the mean scores of the bilingual and the monolingual subjects on the Stanford-Binet and the WISC Full Scale and on the WISC Verbal and Performance Scales in favor of the monolinguals. The

differences in mean scores achieved by the two groups on the Goodenough and the Progressive Matrices were not statistically significant. In total, the monolingual subjects surpassed the bilingual subjects on the verbal tests of intelligence and one of the three performance measures.

Spoerl (1943) conducted an investigation in order to determine the effect of bilingualism upon the measurement of intelligence and upon the academic achievement of selected groups of college students at the American International College. A "survey group" of 69 bilinguals and an "intensive study group" of 32 bilinguals were used. The intensive study group was matched with corresponding numbers of monolingual students with respect to sex, age, and intelligence, as measured by the Henmon-Nelson Test of Mental Ability. The survey group was matched only on the Henmon-Nelson test. In addition, the intensive study group was matched with the control group on the basis of socioeconomic status. To be admitted to the bilingual group, a subject had to have learned the two languages at or before entrance into the first grade of elementary school.

The experimental group subjects were then given the 1937 revision of the Stanford-Binet Intelligence Test, Form L, and the Purdue Placement Test in English. Spoerl found no significant differences between the performances of the bilinguals and monolinguals on the Stanford-Binet. The bilingual subjects received higher scores than the monolingual subjects on the Purdue Placement Test. A review of class records indicated, however, that the bilingual subjects used in this study had done consistently better academic work than the monolingual

subjects. Thus, one would have predicted higher scores for this group on the Purdue Placement Test. Since the groups were matched for ability on the Henmon-Nelson, it might also have been expected that the differences in performance on the Stanford-Binet would not be significant.

Lester's (1923) study was concerned with the results of administering performance tests to foreign-born children and with the advisability of using such tests in the selection of children in school systems. A retarded group of 26 first grade children, 18 of whom came from homes where Polish was spoken always and eight of whom were English-speaking monolingual children were the subjects of the investigation.

The 1916 revision of the Stanford-Binet Scale and several performance tests were administered to all the subjects. The performance tests included the Seguin Form Board, the Mare and Foal, the Knox Cube, the Healy Form Board A, the Kohs Block Design, and the Porteus Maze. These performance tests were administered as group tests and since Lester found considerable difficulty in administering the mazes to the foreign group, she questioned the results.

Important for this review, the average Stanford-Binet IQ for the foreign group was 80.5 and that for the English-speaking groups was 81, a nonsignificant difference. On most performance tests also, the bilingual and monolingual group scores did not differ significantly. The small and unequal number of subjects in each language group and the administration of the performance tests as group tests might serve to invalidate the results of this investigation.

Pintner and Keller (1922) divided children in primary grades from three schools into an English-speaking group and a foreign group. Language classifications were based upon parents' nationalities. There were 367 children in the English-speaking group and 674 in the foreign group. All subjects were given a special revision of the 1916 Stanford-Binet Intelligence Test. The average IQ of the English-speaking only group was found to be 92, while the foreign group was 84. The authors concluded that bilingual children were penalized with lower scores on revisions of the Stanford-Binet. They felt that these children would suffer a serious handicap if classified with English-speaking children on the basis of MA's achieved only on the Stanford-Binet.

Kittell (1963) compared the language and nonlanguage intelligence test scores of a selected group of 83 children from one elementary school in Berkeley, California. There were 42 children in the bilingual group and 41 children in the monolingual control group. The bilingual group included all the children in the upper half of the third grade whose parents stated that a foreign language was spoken at home instead of, or in addition to, English. The control group consisted of children in the third grade whose parents did not indicate that a language other than English was spoken at home.

The California Short Form Test of Mental Maturity, 1953 S Form, and the California Reading Test, Primary Form AA, were administered during the second month of the child's attendance in the upper half of the third grade. The results indicated that the control group's mean language MA was significantly greater than that of the bilingual group,

but there was no significant difference in the mean reading scores of the two groups on the reading test.

The following studies concerned the testing of bilingual (Spanish/English) subjects on an individual intelligence test. Several of these studies tested subjects in both English and Spanish.

Anastasi and de Jesus (1953) obtained measures of language development and Goodenough Draw-A-Man IQ's for 25 bilingual Puerto Rican boys and 25 bilingual Puerto Rican girls who attended day nurseries in New York City. All of the subjects were within six months of their fifth birthdays and their scores were compared with those of 50 white and 50 Negro children.

The Draw-A-Man test results showed bilingual children to be penalized. The mean IQ's on the Goodenough Draw-A-Man Test for the Puerto Ricans were 88.88 for the boys, 102.52 for the girls, and 95.70 for the entire Puerto Rican group, whereas the mean IQ's for the white subjects were 99.00 for the boys, 105.64 for the girls, and 102.32 for the entire white group. The mean IQ's of the Negro subjects were 96.88 for the boys, 103.28 for the girls, and 100.08 for the entire Negro group. It is important to note also that the educational and occupational levels of the parents of the Puerto Rican children were markedly lower than those of the Negro and white parents.

Anastasi and Cordova (1953) administered the Cattell Culture-Free Test, Forms 2A and 2B, to 176 bilingual boys and girls in grades 6, 7, and 8 whose ages ranged from 11 to 15 years. The subjects attended a parochial school in Spanish Harlem of New York City. Half the group received the test instructions of Form A of the test in

English during the first testing session, and this same half received the test instructions of Form B of the test in Spanish during the second testing session. The order of languages was reversed for the other half of the group.

The rating scale to determine the bilingualism of the subjects was adopted from the Hoffman Bilingual Schedule. Nearly all of the parents of the subjects came from Puerto Rico, while nearly half the subjects were born in Puerto Rico and the remaining number in New York City. Anastasi and Cordova found significantly marked improvement of test scores from the first to the second testing session regardless of the order of languages employed. The authors found subjects did just as poorly when directions were in English as when they were in Spanish.

The median standard score IQ of the group as a whole was 70, with an SD of 24 points. This IQ was 1.25 standard deviations below the test norms reported by Cattell in the test manual. The authors attributed this low score to the low socioeconomic status of the parents, to the bilingualism of the subjects which made them deficient in both English and Spanish, to their extreme lack of test sophistication, and to their poor adjustment to the school situation.

Roca (1955) did a study involving the translation and adaptation of three intelligence tests for the use in the public schools of Puerto Rico. The tests were the Wechsler Intelligence Scale for Children, Stanford-Binet, Revised, Form L, and the Goodenough Intelligence Test. Roca says only that the children were usually bilingual. His population did score lower on all three measures than the standardization population. With the translation and adaptation of the Wechsler, the

average Puerto Rican child scored 12 IQ points less than the average American child.

For the Stanford-Binet, a total of 889 subjects were used in three studies. After all changes were made, the scale was then used on 104 new subjects and norms established for Puerto Rico. Comparing these norms with American norms, it was found that the Puerto Ricans scored lower than American monolingual children.

Puerto Rican norms were also established for the Goodenough Intelligence Test established by testing 1,767 children. Comparing these norms with American norms, it was again found that the Puerto Rican children scored lower than monolingual children. Roca suggests that there will always be cultural differences on intelligence tests and that one should consider whatever average is obtained (for the bilingual group) as equivalent to an IQ of 100.

Altus (1953) attempted to compare the intelligence test patterning of a selected portion of the Mexican-descent bilinguals to that of monolingual English-speaking children using the Wechsler Intelligence Scale for Children. The samples were matched on age, sex, and performance IQ. The children in this study were referred to the Guidance Department of the Santa Barbara County School System for preliminary screening for special class placement.

The classification of bilingualism was accepted if the cumulative record cards indicated that Spanish was spoken exclusively at home or that both Spanish and English were spoken at home. The monolingual group came from school referrals also, but the primary reason for referral was maladjustment. All children were tested in English only.

Differences between the two language groups in IQ's on the Verbal Scale of the WISC averaged 17 points in favor of the monolinguals. The bilinguals achieved a mean Verbal IQ of 72.02, whereas the monolinguals achieved a mean Verbal IQ of 88.98. All the verbal subtests showed statistically significant differences at the one percent level of confidence or better between the two groups. There was no significant difference between the monolinguals and bilinguals on the performance scale (they were matched on performance IQ).

Christiansen and Livermore (1970) compared the performance of 92 teenage lower- and middle-class Anglo-American subjects with lower- and middle-class Spanish-American subjects on the Wechsler Intelligence Scale for Children. A child was considered Spanish-American if both parents had Spanish surnames. Lower- and middle-class determination was based on fathers' occupation. Comparisons were made on the following: (a) the full scale scores in English, (b) the Verbal Scale IQ score, (c) the Performance Scale IQ score, (d) the intelligence factors of verbal comprehension, freedom from distractability, perceptual organization, and relevance.

Middle-class children in both ethnic groups scored significantly higher than lower-class children on each of the WISC measures examined in this study. Nonverbal ability and ability in perceptual organization was not related to ethnic origin. Presumably social class was a more important factor than being Spanish-American (often bilingual) or Anglo-American.

Hickey (1972) measured intelligence and verbal learning ability among Mexican-American preschoolers using the Peabody Picture Vocabulary

Test. Hickey deleted verbal nouns which he felt caused semantic confusion. Subjects were randomly drawn from a population of approximately 2,000 Head Start students in Los Angeles, California. In the item analysis phase of the study, two equal groups of 100 subjects were identified as English monolingual (Group A) and Spanish/English bilingual (Group B). In a second phase, two groups of 30 subjects not identified for Group A or B were chosen.

When the PPVT was administered without alteration, there was a clear difference between Groups A and B. When the modified version of the PPVT was administered, there were no differences between the monolingual and bilingual groups (Group C and D). Hickey concludes that bilingualism appeared to be much less analogous to other learning handicaps than has previously been indicated.

Palmer and Gaffney (1972) preselected subjects for their study from the entire fifth grade population of the Nogales Elementary School, randomly selecting one-half. The 150 children involved were given the WISC in English. After one year, a random sample of 30 from the initial 150 were chosen for testing with the Spanish translation of the WISC developed at The University of Arizona. The scaled scores for the Full Scale, Verbal, Performance, and the ten subtests for both Spanish and English test administrations were used. The authors found no significant differences between the English and Spanish testings on any of the subtests.

Levandowski (1975) attempted to determine if there was a significant difference in intelligence test scores of bilingual students

on an English version of the intelligence test as compared to the scores of the Spanish translation of the test. The author first administered the Hoffman Bilingual Schedule, then the Slosson Intelligence Test was administered to 18 subjects from Spanish-speaking home environments. The Slosson was translated into Spanish for the purpose of this study. The split half method was used to administer the Spanish and English versions to all children.

The mean scores on the Spanish test were 107, while the mean for the English was 89. The median for the Spanish and English tests were 105 and 94.5 respectively. Levandowski (1975, p. 50) concludes that bilinguals are penalized by these tests when administered in English. "On the basis of intelligence test scores of a test that is biased in construction, biased in the norms for standardization, and administration in a language other than the mother tongue of the child, bilingual students are classified as 'slow learners' or 'underachievers'."

Discussion

The majority of these representative studies seem to indicate that bilingual children are penalized when administered individual verbal intelligence tests in English. Indications are that bilingual children are not penalized when administered individual nonverbal intelligence tests in English. Before drawing any conclusions, however, these studies must be collectively reviewed.

The majority of studies tested children in English only. That is, they did not test the same children in two languages and then compare groups. These studies were Pintner and Keller (1922), Lester

(1923), Seidl (1937), Spoerl (1943), Darcy (1946), Altus (1953), Anastasia and de Jesus (1953), Roca (1955), Levinson (1959), Kittell (1963), and Christiansen and Livermore (1970). One cannot help but notice that all of these studies would lead to the indicated conclusion that bilingual children are penalized on individual verbal intelligence tests administered in English and not penalized on nonverbal individual intelligence tests administered in English. However, one gets a different picture when the studies that tested children in two languages are reviewed.

The studies that tested children in two languages are: Anastasia and Cordova (1953), Hickey (1972), Palmer and Gaffney (1972), and Levandowski (1975). Three of these four studies found that bilingual children are not penalized on individually administered intelligence tests. Anastasia and Cordova (1953) used the Cattell Culture-Free Test and used the split half method to test children on the English and Spanish version. No significant differences between the two versions were found, and it was hypothesized that low scores were due to the socioeconomic level of the parents. Hickey (1972) used the Peabody Picture Vocabulary Test and found no significant differences between children tested in English or Spanish when verbal nouns were removed from the test. Palmer and Gaffney (1972) randomly selected fifth graders in a border town between Mexico and the United States and found no significant differences on the WISC administered in English or Spanish.

Levandowski (1975) administered the Slosson to bilingual children in Spanish and English. He found that the children were penalized

when tested in English. Levandowski felt this was due to the biased construction of the test and biased norms for standardization.

One cannot ignore the fact that three out of the four studies reviewed which tested children in both Spanish and English leave a serious question as to whether or not bilingual children need to be tested in their primary language because they are penalized in English. The following conclusion is drawn from the above evidence: It is not yet possible to answer the question of whether bilingual children are penalized on individually administered verbal and nonverbal intelligence tests when tested in English.

Summary

This chapter reviewed the research concerned with testing the bilingual child. Studies using group intelligence tests were reviewed. Because of inadequacies in group tests, a review of these studies does not help to answer the question of whether bilingual children perform better on tests given in English or Spanish. The research using individually administered intelligence tests is inconclusive at best. To answer this question, bilingual children should be administered an instrument standardized in Spanish and an English translation and the scores compared.

CHAPTER 3

QUESTIONS, METHODS, AND PROCEDURES

Statement of the Problem

As can be seen from the review of literature in Chapter 2, there is a great deal of controversy dealing with the use of tests standardized on English-speaking populations and used on children from Spanish-speaking backgrounds. Zirkel (1972) points out that translating standardized achievement tests from English to Spanish points to, but does not provide, the way to more equitable opportunities for bilingual children. The development of tests appropriate to Spanish children is more than simply translating existing tests. Many researchers have found that bilingual children do less well than monolingual children, but this is usually when an English test is translated into Spanish.

No one has yet looked at the question of whether bilingual children test higher on a test first standardized in Spanish and then translated into English. In short, will bilingual children do better on the Spanish standardized version or on the English translation of the ITPA? This study attempts to answer this question by examining the following specific questions.

1. Is there a significant difference between scores on the Spanish ITPA and its English translation within a kindergarten and second grade sample?

2. Is there a significant difference between auditory and visual channel scores of bilingual children on the Spanish ITPA and its English translation?
3. Is there a significant testing order effect for the kindergarten and second grade groups?
4. Is there a difference in mean deviation scores of individual subjects on the Spanish standardized ITPA and its English translation?
5. How do scores of subjects in this study's sample compare to scores of monolingual Latin American children and to scores of bilingual children in Tucson, San Diego, and Miami on the Spanish ITPA?

Limitations of the Study

This study is limited in that it deals with an instrument standardized in Spanish and its English translation. The results should be of great interest to those concerned with testing of bilingual children and help to answer questions arising from recent court decisions requiring that children be tested in their primary language. However, generalizability of the results is limited due to the particular idiosyncrasies of any test and the sample.

Subject Selection and Population

The subjects for this study were selected from the total population of children in the kindergarten and second grade bilingual program in the Hayward Unified School District in Hayward, California.

Hayward Unified School District has a total student population of 20,015. It is in suburban San Francisco and has one of the largest bilingual programs in the country. The bilingual program is voluntary. There is a total of 132 children in the kindergarten bilingual program and 165 children in the second grade bilingual program.

To select the subjects for the study, the master computer list maintained by the district was utilized to gain the names of children involved in the bilingual program who were orally fluent in both Spanish and English. The bilingual program assistants (who supervise the bilingual teachers), and who are familiar with the children, teachers, and schools, selected those children that they felt were "most" bilingual. All of the kindergartners and some of the second graders had been administered the Bilingual Syntax Measure yielding a score in both languages. The San Diego Observation Assessment Instrument had been administered to many of the students. The following criteria were used in selecting "bilingual children."

1. All children were in the bilingual program on a voluntary basis.
2. Teacher and program assistant both agreed that the child had developed expressive and receptive use of both Spanish and English.
3. The children and parents reported that Spanish is spoken in the home of the child.

Of the total population of children in the bilingual program, 28 children met this criteria at the second grade level and 26 at the kindergarten level. The children came from eight different elementary

schools in the Hayward district. The schools were Bowman, Burbank, Cherryland, Lorin Eden, Muir, Palma Cela, Park, and Ruus. The bilingual program utilizes both English and Spanish as a means of instruction. Reading is first taught in Spanish and later in English.

The population of Hayward is basically middle income and includes lower income families. The racial makeup of the Hayward area is Anglo, 12,500; Hispanic-American, 4,300; Black, 1,600; Asian, 900; Filipino, 500; and American Indian, 200. More detailed descriptive information is provided on the subjects in Chapter 4 dealing with results.

Descriptive Analysis of Subjects

Table 1 gives the absolute frequency and relative frequency for sex of subjects in kindergarten, second grade, and combined. As can be seen from Table 1, slightly more than half the kindergarten subjects and second grade subjects were female. The total number of males in the study was 24 or 44% of the group, and the total number of females in the study was 30 or 56%.

Ages of Subjects. The kindergarten subjects were all five and six years old, whereas the second grade subjects were ages seven, eight, and nine at the time the tests were administered.

Subjects Who Attended Preschool. About one-half of the kindergarten and one-half of the second grade subjects had attended preschool.

Table 1. Sex of subjects

	Absolute Frequency	Relative Frequency (%)
Kindergarten		
Males	11	42%
Females	15	58%
Total	N = 26	100%
Second Grade		
Males	13	46%
Females	15	54%
Total	N = 28	100%
Kindergarten and Second Grade Combined		
Males	24	44%
Females	30	56%
Total	N = 54	100%

Handicaps of Subjects. None of the subjects were reported by their teachers to have any hearing, vision, or other physical handicaps. All 54 subjects were in the bilingual program and none of the subjects had special education class placement or services.

Grade of Subjects. Twenty-six subjects were in kindergarten and 28 subjects were in the second grade.

Subjects' Parental Birthplace. Table 2 shows the birthplace of the subjects' mothers and fathers. Approximately 80 percent of the mothers of kindergarten subjects and second grade subjects were born in Spanish-speaking countries. Approximately 85 percent of the fathers of kindergarten subjects and second grade subjects were born in Spanish-speaking countries. The remaining parents were born in the United States.

Subjects' Birthplace. An overwhelming majority--80 percent--of the subjects were born in California. The remaining 20 percent of the subjects minus one were born in Mexico.

Socioeconomic Status of Subjects' Families. Hayward is a middle- and lower-socioeconomic class city by most standards. Subjects' families were considered low socioeconomic status if they qualified for free meals at school, and middle-class if they did not. Teacher judgment was also asked. Approximately 28 percent of the subjects were from middle socioeconomic families and 72 percent were from lower socioeconomic status families.

Table 2. Birthplace of subjects' parents

	Absolute Frequency	Relative Frequency
Kindergarten		
Mothers		
Spanish-speaking countries	21	80%
California	2	12%
Other states in U.S.	3	8%
Total	N = 26	100%
Fathers		
Spanish-speaking countries	22	85%
California	2	7.5%
Other states in U.S.	2	7.5%
Total	N = 26	100%
Second Grade		
Mothers		
Spanish-speaking countries	22	78%
California	5	18%
Other states in U.S.	1	4%
Total	N = 28	100%
Fathers		
Spanish-speaking countries	24	86%
California	2	7%
Other states in U.S.	2	7%
Total	N = 28	100%

District and Schools of Subjects. All 54 subjects attended eight different elementary schools in Hayward Unified School District.

Ethnic Membership. The families of all 54 subjects used in this study considered themselves to be of Spanish-American culture.

Language Spoken at Home. All 54 subjects reported that Spanish was the language spoken in their homes.

Teachers' Bilingual Judgment. The teachers rated 45 of the 54 subjects as bilingual. Teachers said they felt the other nine children spoke Spanish better than English.

Program Assistants' Bilingual Judgment. Program assistants help the teachers with the program development for children in the bilingual program and know the children well. The program assistant rated all 54 subjects as bilingual.

Testing Procedure

All subjects were tested after obtaining permission slips from their parents. All subjects were tested on the Spanish ITPA and its English translation. The split-half method was used. Half of the children at each grade level were tested first in Spanish and then tested within four days in English. The other half of the subjects were tested in English first and administered the Spanish test within four days of the English administration. All testing was done under quiet conditions in the school the child attended, after first gaining his/her agreement.

Any talking before testing was done by the examiner in the language in which the child was to be tested. All of the tests were administered by the bilingual investigator trained in the administration of the ITPA.

The Instrument

The Illinois Test of Psycholinguistic Abilities. The ITPA is based on a model of language acquisition devised by Osgood (1957) and extended by McCarthy and Kirk (1961) and by Kirk, McCarthy, and Kirk (1968) in their revised edition of the test. Osgood (1957) has devised a comprehensive psycholinguistic model from which a diagnostic test can be developed as a basis for the planning of remedial programs. The Osgood model has been modified by McCarthy and Kirk (1961) by including short-term memory components. In this respect, the abilities at the automatic level resemble the perceptual concepts in the Wepman (1958) model.

The clinical model of the ITPA postulates two levels: the representational level and the automatic level. A description of these levels follows.

Representational Level. Three general types of functions are tested at the representational, or symbolic, level--the receptive process, the organizing process, and the expressive process.

The receptive process may be defined as the process whereby children are able to comprehend symbols. Receptive process tests include the following:

1. Auditory reception test assesses ability to derive meaning from the spoken word. Short direct questions requiring only yes or no answers are included.
2. Visual reception test assesses ability to derive meaning from visual symbols. The child is shown a stimulus picture for three seconds and is told, "See this." Then a page of pictures, one of which is conceptually similar to the stimulus picture, is shown. The subject is asked to point to the correct picture.

The organizing process may be defined as the ability to organize or relate symbols in a meaningful way.

1. Auditory-vocal association test assesses ability to relate concepts presented orally. The subject is asked to complete sentences that are analogies. "You sit in a chair; you sleep in a _____."
2. Visual-motor association test uses picture association. A single stimulus picture is surrounded by four optional pictures. The child points to the optional picture that goes with the stimulus picture.

The expressive process is the ability to transmit an idea through words or gestures.

1. Verbal expression test involves asking the child to tell about a number of different objects.
2. Manual expression test assesses the ability to use manual expression in response to being asked, "Show me what you do with" questions. The child pantomimes the response.

Automatic Level. Two types of functions are measured at the automatic or nonmeaningful level: closure and memory. Closure is the ability to fill in missing parts in an incomplete picture or expression.

1. Auditory closure test assesses the child's ability to fill in the missing parts of words as might be heard in a faulty phone connection.
2. Visual closure test assesses the ability to identify common objects from an incomplete picture.

Sequential memory is the ability to reproduce a sequence of stimuli from memory.

1. Auditory sequential memory test assesses ability to reproduce sequences of digits that have been presented auditorally.
2. Visual sequential memory test assesses ability to reproduce sequences of visual figures from memory. Nonmeaningful figures are used.

There have been many demands in research in recent years for a Spanish ITPA. Recently, a project was undertaken to develop the Spanish ITPA--not a translation but a new test looking at the same psychological processes but standardized on Spanish-speaking children.

For developing norms, children were tested using the experimental version in New York, Miami, San Diego, Tucson, Puerto Rico, Venezuela, Columbia, Mexico, Chile, and Peru. After testing was completed in all areas, items of dialectal variations or poor discriminatory power were eliminated.

The Spanish ITPA consists of the basic subtests on the English version, but many of the procedures and practically all of the items are different. In short, it is a new diagnostic instrument standardized in Spanish. The Spanish ITPA must be administered individually by a Spanish-speaking examiner. It is not known at this time, however, if a test produced in Spanish will yield higher scores for bilingual children than the same test translated into English.

Test Scoring

All tests were scored three times to insure accuracy. The examiner first scored all of the protocols. Next a computer programmer experienced in working with data rescored all of the protocols, rechecking also all of the stanines. Finally, a third person checked both of these scoring procedures. As a further added extra measure, the possible score ranges were written into the computer program for the statistical analysis to insure no gross errors were included but instead would be rejected by the computer. The following 10 subtests were given in Spanish: Associacion Visual, Associacion Auditiva, Comprension Visual, Comprension Auditiva, Fluidez Lexica, Expresion Motoro, Integracion Auditiva, Integracion Visual, Memoria Secuencial Auditiva, Memoria Secuencial Visomotora.

In addition to scores for each of these subtests, a mean score for the total test was obtained, a mean for the six tests dealing with the representational level, a mean for the four tests dealing with the automatic level, a mean for the five tests dealing with the auditory/vocal channel and a mean for the five tests dealing with the visuomotor channel.

This same procedure was used in the English translation with the exception that Integracion Auditiva (Auditory Integration) could not be translated into English so for the English administration of the test, Auditory Closure from the English ITPA was administered.

Statistical Treatment of the Data

Analysis of Data

The general question is, Do bilingual children do better on the Spanish version of the ITPA or the English translation? The five specific questions addressed and the statistical treatment of these questions is described below.

Question 1: Is there a significant difference between scores on the Spanish ITPA and its English translation within the kindergarten and second grade samples? A two-way analysis of variance using repeated measures was conducted to compare the raw scores of the Spanish version subtests and the English translation.

	A English Translation	B Spanish Version
Kindergartners (N = 26)	26	26
Second Graders (N = 28)	28	28
Total N = 54		

Question 2: Is there a significant difference between auditory and visual channel scores of bilingual children on the Spanish ITPA and its English translation? A two-way analysis of variance using repeated measures was conducted comparing the auditory-vocal channel scores to the visual motor channel scores.

Question 3: Is there a significant testing order effect for the kindergarten and second grade groups? A two-way analysis of variance using repeated measures was conducted looking at order effects. The following diagram presents the number of subjects in each cell.

	Testing Order Effects for Kindergarten Subjects	
	First Administration	Second Administration
Groups		
Spanish Test	13	13
English Test	13	13

This same design was used for second grade subjects with N = 14 in each cell.

Question 4: Is there a difference in mean deviation scores of individual subjects on the Spanish standardized ITPA and its English translation? Total deviation scores were calculated by summing the deviations of the scaled scores of each subtest from the child's own mean without regard to sign and dividing by 10 (using the Kirk and Elkins' 1974 formula). Those subjects falling within the lowest three percent of the population are considered to have severe learning disabilities.

Question 5: How do scores of subjects in this study's sample compare to scores of monolingual Latin American children and to scores of bilingual children in Tucson, San Diego, and Miami on the Spanish ITPA? A simple t test was conducted comparing the means of the various groups.

CHAPTER 4

RESULTS

The purpose of this study was to determine whether bilingual Hispanic children do better on a test standardized in Spanish or its English translation. The results of this study will attempt to answer the following questions:

1. Is there a significant difference between scores of the Spanish ITPA and its English translation within the kindergarten and second grade samples?
2. Is there a significant difference between auditory and visual channel scores of bilingual children on the Spanish ITPA and its English translation?
3. Is there a significant testing order effect for the kindergarten and second grade groups?
4. Is there a difference in mean deviation scores of individual subjects on the Spanish standardized ITPA and its English translation?
5. How do scores of subjects in this study's sample compare to scores of monolingual Latin American children and to scores of bilingual children in Tucson, San Diego, and Miami on the Spanish standardized ITPA?

Comparison of Spanish ITPA and English Translation

Results of the Spanish vs. English ITPA

This section will present the results of the differences between scores on the Spanish ITPA and its English translation for (a) the kindergarten subjects and (b) the second grade subjects.

Is there a significant difference between raw scores on the Spanish ITPA and its English translation within the kindergarten sample? Table 3 presents the ANOVA results for repeated measures for the raw scores of the subtests of the Spanish version of the ITPA and its English translation for 26 kindergarten children. It will be seen from Table 3 that the two-way ANOVA for repeated measures shows:

1. There were significant differences in two subtests (Auditory Comprehension $p < .01$, and Auditory Sequential Memory $p < .01$) in favor of the English translation.
2. There were no significant differences in the other eight subtests between the two administrations of the test.
3. There were no significant differences on the Spanish versus English translation in the auditory-vocal channel or in the visual motor channel.
4. There were no significant differences in the means on the Spanish or English versions at the representational level or the automatic level.
5. There were no significant differences in the total test means between the Spanish standardized test and the English translation for kindergarten subjects.

Table 3. Raw score means, standard deviations and significance of difference between Spanish and English translation of ITPA administered to kindergarten subjects (N = 26)

Subtests	Spanish Mean	Spanish SD	English Mean	English SD	F Ratio	p
Visual Association	12.73	3.39	12.00	3.01	2.2	NS
Auditory Association	8.77	5.98	8.12	5.09	.25	NS
Visual Comprehension	14.46	4.08	13.77	3.73	.88	NS
Auditory Comprehension	7.0	3.86	9.89	3.14	10.97	<.01
Verbal Fluency	25.15	10.18	26.40	9.54	1.03	NS
Motor Expression	14.81	3.25	14.88	3.13	.03	NS
Auditory Integration	10.42	3.99	9.43	3.46	1.08	NS
Visual Integration	33.23	6.95	35.08	6.61	1.40	NS
Auditory Sequential Memory	6.15	1.95	8.27	2.63	34.36	<.01
Visual Sequential Memory	6.57	2.83	7.00	2.80	.51	NS
Representational Level Mean	70.58	16.96	72.71	17.93	.50	NS
Automatic Level Mean	51.45	9.66	54.52	9.48	2.22	NS
Auditory Vocal Channel Mean	52.58	17.34	55.54	17.69	.89	NS
Visual Motor Channel Mean	76.55	10.28	77.13	10.00	.09	NS
Total Test Means	133.39	24.15	138.59	25.17	1.45	NS

It is obvious from the above results that when tested on the Spanish and English versions of the ITPA the bilingual children in this sample of kindergarten children differ in performance on only two subtests. On both of these subtests the kindergarten subjects did better on the English version of the test.

Table 4 presents the ANOVA results for repeated measures for the raw scores of the subtests of the Spanish version of the ITPA and its English translation for 28 second grade subjects. It will be seen from Table 4 that the ANOVA for repeated measures shows:

1. In contrast to the kindergarten group, the second grade group shows significant differences in favor of the English translation of the test in (a) four auditory-vocal subtests: Auditory Association ($p < .05$), Verbal Fluency ($p < .05$), Auditory Comprehension ($p < .01$), and Auditory Sequential Memory ($p < .01$); (b) the representational level; (c) the auditory vocal channel, and (d) the total test mean.
2. There were no significant differences in (a) Visual Association, Visual Comprehension, Motor Expression, Auditory Integration, Visual Integration, and Visual Sequential Memory; (b) automatic level mean; and (c) visual motor channel.

On the auditory-vocal subtests for the second grade subjects, it appears that several years of schooling influence children's performance on the English translation of the test.

Table 4. Raw score means, standard deviations and significance of difference between Spanish and English translation of ITPA administered to second grade subjects (N = 28)

Subtests	Spanish Mean	Spanish SD	English Mean	English SD	F Ratio	p
Visual Association	16.18	2.85	16.29	2.93	.04	NS
Auditory Association	12.72	8.28	17.93	7.79	4.70	<.05
Visual Comprehension	16.86	2.97	16.86	2.89	0	NS
Auditory Comprehension	12.85	5.97	15.71	6.20	4.05	<.05
Verbal Fluency	32.67	8.99	42.64	12.97	15.28	<.01
Motor Expression	17.10	4.00	16.25	3.74	1.70	NS
Auditory Integration	14.14	5.49	15.61	4.51	1.25	NS
Visual Integration	44.32	4.79	44.10	6.34	.04	NS
Auditory Sequential Memory	8.18	2.31	10.68	3.13	16.40	<.01
Visual Sequential Memory	14.21	2.75	14.21	2.78	0	NS
Representational Level Mean	99.14	19.31	112.14	25.60	9.74	<.01
Automatic Level Mean	70.20	9.36	73.95	10.56	2.25	NS
Auditory Vocal Channel Mean	74.03	22.41	94.03	26.41	8.94	<.01
Visual Motor Channel Mean	97.31	8.28	96.34	11.82	.23	NS
Total Test Means	176.47	26.87	197.49	35.88	6.59	<.05

Results of Auditory vs. Visual Channel Scores

Differences between languages would be expected to occur in the auditory-vocal channel of communication and not in the visual motor channel. To shed some light on this question, the data were analyzed to determine whether there was a significant difference between the auditory-vocal and visual motor channels of the Spanish standardized test and its English translation for both the kindergarten and second grade groups.

Is there a significant difference between stanine scores of bilingual children on the auditory and visual channels of the Spanish ITPA and its English translation? Tentative norms were provided for the standardized Spanish edition of the ITPA. Although these norms are tentatively applicable to Hispanic children, in order to compare inter-channel differences across both visual and auditory channels, it is necessary to translate the data available on individual subtests into standard scores or stanines. For this comparison stanine scores were utilized.

Table 5 presents the stanine means of the auditory-vocal channel, visual motor channel, standard deviations, F ratio, and significance of the difference between the Spanish ITPA and the English translation administered to kindergarten and second grade subjects. Following are the conclusions drawn from Table 5.

It will be noted from this table that kindergarten and second grade subjects scored significantly poorer on the auditory-vocal channel than they did on the visual motor channel on both the Spanish and English tests. These differences were significant at the .001 level.

Table 5. Stanine means of auditory-vocal channel, visual motor channel, standard deviations, F ratios, and significance of difference between Spanish standardized ITPA and English translation administered to kindergarten and second grade subjects (N = 54)

	Auditory-Vocal Channel		Visual Motor Channel			
	Mean	SD	Mean	SD	F Ratio	p
Kindergarten (N = 26)						
Spanish	3.05	.99	4.89	1.00	70.41	<.001
English	3.40	1.13	4.90	1.15	45.74	<.001
Second Grade (N = 28)						
Spanish	3.15	1.27	5.46	1.11	77.47	<.001
English	4.11	1.59	5.43	1.31	24.22	<.001

It would appear that with the bilingual children of this sample there is a significant difference between performance on tests relying on the auditory-vocal channel compared to the visual motor channel. The superiority of this group in the visual motor channel indicates that one must be cautious in administering auditory-vocal tests to bilingual children. Table 5 supports the supposition that bilingual subjects are penalized on an auditory-vocal test compared to a visual motor test whether the test is in English or Spanish.

Effect of Testing Order

Although alternating the testing order should control for practice effects, the data were analyzed to determine whether testing order effects resulted using this design.

Is there a significant testing order effect for the kindergarten or second grade groups? Table 6 presents the ANOVA results for repeated measures for raw scores of the 26 kindergarten subjects examining differences between means of the first and second administration of the ITPA ignoring which language the subject was tested in first. It will be seen from Table 6 that the ANOVA for repeated measures shows there were significant testing order effects on three subtests (Motor Expression $p < .05$, Visual Integration $p < .01$, and Visual Sequential Memory $p < .01$) in favor of the second administration. This would indicate a practice effect on these subtests.

Table 7 presents the ANOVA results for repeated measures for raw scores of 28 second grade subjects. Differences between means on the first and second administration of the ITPA were examined ignoring which language the subject was tested in first. It will be seen from Table 7 that the ANOVA for repeated measures shows there were significant testing order effects on five subtests (Visual Comprehension p .05, Verbal Fluency $p < .05$, Auditory Integration $p < .05$, Visual Integration $p < .01$, and Visual Sequential Memory $p < .01$) in favor of the second administration.

It is not surprising that testing order effects or practice effects are significant for some subtests at the kindergarten and second grade level. One would expect that since the subjects are bilingual and

Table 6. Kindergarten raw score means for each subtest by testing order only (N = 26)

	$\overline{X}$ First Administration	$\overline{X}$ Second Administration	p
Visual Association	11.92	12.80	NS
Auditory Association	7.27	9.61	NS
Visual Comprehension	13.62	14.61	NS
Auditory Comprehension	8.15	8.73	NS
Verbal Fluency	25.42	26.19	NS
Motor Expression	14.34	15.34	<.05
Auditory Integration	9.38	10.46	NS
Visual Integration	31.46	36.84	<.01
Auditory Sequential Memory	7.03	7.38	NS
Visual Sequential Memory	6.04	7.54	<.01

Table 7. Second grade raw score means for each subtest by testing order only (N = 28)

	$\bar{X}$ First Administration	$\bar{X}$ Second Administration	p
Visual Association	15.75	16.71	NS
Auditory Association	13.79	16.85	NS
Visual Comprehension	16.29	17.43	<.05
Auditory Comprehension	13.00	15.57	NS
Verbal Fluency	35.07	40.25	<.05
Motor Expression	16.18	17.18	NS
Auditory Integration	13.64	16.10	<.05
Visual Integration	41.82	46.60	<.01
Auditory Sequential Memory	9.21	9.64	NS
Visual Sequential Memory	13.25	15.17	<.01

being presented the same material twice, only in different languages, that there would be some practice effects. Importantly, because of the split-half design, the testing order effects that exist in no way invalidate the results of question one which looked at performance in Spanish versus performance on the English translation.

Differences in Individual Deviation Scores

One of the questions that can be raised is whether or not greater deviations are found among bilingual children on a Spanish standardized test or an English translation. Learning disabilities has been defined as constituting a discrepancy between developmental abilities. For this reason the data were analyzed to answer the following question.

Is there a difference in mean deviation scores of individual subjects on the Spanish standardized ITPA and its English translation? The norms on the English edition of the ITPA indicate that an average deviation of 6.7 and above defines the lowest three percent of the population (Paraskevopoulos and Kirk, 1968). Using the same formula and extrapolating from the English norms, a total deviation score of 18 and above on the Spanish edition would define the lowest three percent of the population.

The total deviation scores were calculated by summing the deviations of the scaled scores of each subtest from the child's own mean, without regard for sign, and dividing by 10. The following formula was used:

$$\text{Mean Deviation} = \frac{\Sigma \left| X_i - \bar{X} \right|}{10}$$

where X_i = scaled score of subtest

$\bar{X}$ = mean scaled score (Kirk and Elkins, 1974)

The following conclusions can be drawn from Table 8:

1. At the kindergarten level none of the children exhibit a discrepancy greater than 18 when tested in Spanish. When tested in English, two of the children, #9 and #18, show discrepancies large enough to fall within the lowest three percent of the population and be classified as having severe learning disabilities.
2. By second grade, however, Table 9 indicates that six children show discrepancies of 18 or more when tested in Spanish. Only one child shows such a discrepancy when tested in English.
3. All of the subjects found to meet the criteria of learning disabilities did poorer on the auditory-vocal channel than the visual motor channel. This supports the parametric data presented in question 3 showing a gap by second grade in subjects' performance on auditory-vocal subtests.

When examining individual cases it becomes clear that by second grade there is a depression of auditory-vocal subtest scores and higher scores in the visual motor channel creating wide mean deviations for individual subjects. If this criterion is used (mean deviation score

Table 8. Total test mean deviations for the kindergarten group comparing the Spanish ITPA to the English translation (N = 26)

Subject Number	Mean Deviation Spanish	Mean Deviation English
001	14.80	10.30
002	12.48	11.60
003	5.28	7.36
004	8.80	10.68
005	12.52	12.90
006	15.00	12.80
007	8.60	12.84
008	10.40	10.44
009	8.30	18.10*
010	15.40	11.10
011	11.90	12.60
012	12.00	13.51
013	9.70	7.44
014	13.70	12.88
015	11.20	9.20
016	14.80	10.96
017	10.24	11.30
018	10.80	19.04*
019	7.50	9.36
020	16.28	13.88
021	13.48	16.90
022	16.60	12.08
023	12.70	7.90
024	13.80	13.05
025	14.80	14.64
026	12.70	14.40

*lowest three percent of the population labeled learning disabilities

Table 9. Total test mean deviations for the second grade comparing the Spanish ITPA to the English translation (N = 28)

Subject Number	Mean Deviation Spanish	Mean Deviation English
027	17.52	11.00
028	15.68	14.64
029	12.88	13.20
030	14.30	5.92
031	8.16	12.16
032	20.10*	7.90
033	12.16	17.60
034	18.44*	10.56
035	18.10*	15.55
036	12.32	6.60
037	10.80	14.20
038	12.25	14.72
039	11.40	19.68*
040	13.84	14.42
041	15.80	16.90
042	19.18*	14.30
043	11.04	12.04
044	14.60	10.92
045	15.50	10.70
046	18.20*	12.20
047	14.52	8.00
048	13.28	11.08
049	11.60	6.20
050	15.00	8.82
051	12.40	12.91
052	19.12*	10.52
053	17.46	8.20
054	12.30	13.92

*lowest three percent of the population labeled learning disabilities

<18), children may be identified as learning disabled due to a language problem in Spanish where such a problem does not exist in English. It may result in misclassifying children as learning disabled based on their performance on the Spanish standardized test. If the courts continue to mandate that children be tested in their native language, these children may be incorrectly identified as learning disabled.

Comparison of Bilingual Hispanic Children with Monolingual Hispanic Children

As indicated earlier, the experimental Spanish edition of the ITPA was tentatively standardized on monolingual children in five countries (Columbia, Mexico, Peru, Puerto Rico, and Chile). An attempt was made here to compare the scores of the Hayward group with monolingual Latin-American children. The following question was asked:

How do the scores of subjects in this study compare to scores of monolingual Latin-American children and to scores of bilingual children in Tucson, San Diego, and Miami on the Spanish ITPA? To better understand the sample used in this study, mean raw scores of subjects from this study were compared to mean raw scores of monolingual children from five foreign countries (Columbia, Mexico, Peru, Puerto Rico, and Chile) and of bilingual subjects from Tucson, San Diego, and Miami. Table 10 shows the raw scores, standard deviations and t tests for the Hayward kindergarten subjects compared to these groups. Table 11 shows similar results for the second grade subjects. Several conclusions can be drawn.

Table 10. Raw scores, standard deviations, and t test for Hayward kindergarten subjects compared to monolingual foreign subjects, San Diego-Tucson bilingual subjects, and Miami bilingual subjects tested on the Spanish ITPA

	Hayward Bilingual Kindergarten (N=26)		Monolingual Foreign 5 yr. olds (N=110)		t	San Diego-Tucson Bilingual 5 yr. olds (N=40)		t	Miami Bilingual 5 yr. olds (N=20)		t
	Mean	SD	Mean	SD	Test	Mean	SD	Test	Mean	SD	Test
Visual Association	12.73	3.39	12.51	4.60	NS	12.38	3.78	NS	13.90	2.00	.05
Auditory Association	8.77	5.98	8.36	7.38	NS	8.80	4.61	NS	15.10	5.72	.01
Visual Comprehension	14.46	4.08	12.86	4.58	.05	10.33	4.09	.01	14.30	2.74	NS
Auditory Comprehension	7.00	3.86	13.45	5.30	.01	6.65	5.28	NS	11.05	5.55	.01
Verbal Fluency	25.15	10.18	35.27	10.70	.01	21.50	7.74	.05	30.45	9.63	.05
Motor Expression	14.81	3.25	13.26	3.24	NS	10.75	3.35	.05	11.50	2.91	.05
Auditory Integration	10.42	3.99	11.31	4.57	NS	9.70	3.99	NS	10.15	3.07	NS
Visual Integration	33.23	6.95	30.02	6.21	NS	27.63	6.95	.05	31.70	6.03	NS
Auditory Sequential Memory	6.15	1.95	7.85	2.91	NS	5.10	4.01	NS	7.95	2.63	NS
Visual Sequential Memory	6.57	2.83	7.66	3.38	NS	5.58	2.58	NS	7.50	2.82	NS

Table 11. Raw scores, standard deviations, and t test for Hayward second graders compared to monolingual foreign subjects, San Diego-Tucson bilingual subjects, and Miami bilingual subjects tested on the Spanish ITPA

	Hayward Bilingual 2nd Graders (N=24)		Monolinguals Foreign 7 yr. olds (N=110)		t	San Diego-Tucson Bilingual 7 yr. olds (N=40)		t	Miami Bilingual 7 yr. olds (N=20)		t
	Mean	SD	Mean	SD	Test	Mean	SD	Test	Mean	SD	Test
Visual Association	16.18	2.85	16.74	3.50	NS	16.12	2.33	NS	18.40	2.60	.05
Auditory Association	12.72	8.28	26.60	6.40	.01	15.08	6.49	.05	25.15	8.85	.01
Visual Comprehension	16.86	2.97	16.18	2.98	NS	15.80	2.61	NS	18.00	2.64	NS
Auditory Comprehension	12.85	5.97	18.48	3.74	.01	12.35	5.13	NS	18.65	4.49	.01
Verbal Fluency	32.67	8.99	45.99	13.58	.01	31.40	7.16	NS	39.20	10.37	.01
Motor Expression	17.10	4.00	15.51	3.62	NS	13.83	3.36	.01	14.95	1.73	.01
Auditory Integration	14.14	5.49	14.76	4.17	NS	12.05	4.40	.05	12.35	3.76	.05
Visual Integration	44.32	4.79	38.62	5.73	.05	38.40	6.24	.05	44.50	4.10	NS
Auditory Sequential Memory	8.18	2.31	9.96	2.75	NS	7.15	2.16	NS	10.75	3.80	.01
Visual Sequential Memory	14.21	2.75	12.90	3.14	NS	11.20	5.20	.01	13.95	2.70	NS

The Hayward subjects scored lower than the monolingual subjects on some auditory-vocal tests and higher on some visual motor tests of the Spanish ITPA.

1. The Hayward kindergarten subjects (Table 10) scored significantly lower than the monolingual group on two out of five auditory-vocal tests (Auditory Comprehension and Verbal Fluency). The Hayward second grade subjects (Table 11) also scored significantly lower than the monolingual subjects on Auditory Comprehension and Verbal Fluency, as well as Auditory Association or on three out of five auditory-vocal subtests.
2. The Hayward kindergarten subjects (Table 10) scored significantly higher than the monolingual subjects on Visual Comprehension. Although the difference on Visual Comprehension was not significant at second grade, the Hayward subjects scored significantly higher than the monolingual subjects on Visual Integration at second grade. There were no other significant differences between these groups.

The Hayward subjects scored higher than the combined San Diego-Tucson subjects on several visual motor tests and several auditory-vocal tests.

1. The Hayward kindergarten subjects (Table 10) scored significantly higher than the San Diego-Tucson subjects on Verbal Fluency. The Hayward second grade subjects (Table 11) scored significantly lower than the San Diego-Tucson subjects on one

auditory-vocal test (Auditory Association) and higher on one auditory-vocal test (Auditory Integration).

2. The Hayward kindergarten subjects (Table 10) scored significantly higher than the San Diego-Tucson subjects on three visual motor tests (Visual Comprehension, Motor Expression, and Visual Integration). The Hayward second grade subjects (Table 11) scored significantly higher than the Tucson-San Diego subjects on three visual motor tests (Motor Expression, Visual Integration, and Visual Sequential Memory).

In general the Hayward group scored lower than the Miami subjects on auditory-vocal and visual motor tests.

1. The Hayward kindergarten subjects (Table 10) scored significantly lower than the Miami subjects on three auditory-vocal tests (Auditory Association, Auditory Comprehension, and Verbal Fluency). The Hayward second grade subjects (Table 11) scored significantly lower than the Miami subjects on four auditory-vocal tests (Auditory Association, Auditory Comprehension, Verbal Fluency, and Auditory Sequential Memory). The Hayward second grade subjects (Table 11) scored higher than the Miami subjects on one auditory-vocal test (Auditory Integration).
2. The Hayward kindergarten subjects (Table 10) scored significantly higher than the Miami subjects on one visual motor test (Motor Expression). The Hayward second grade subjects (Table 11) scored lower than the Miami subjects on one visual

visual motor test (Visual Association) and higher on one (Motor Expression).

It appears from these results that the monolingual children in the standardization sample are superior to bilingual children on 90 percent of the auditory-vocal subtests on the Spanish ITPA. This was not true for the Miami subjects whose differences may be due to social class.

Discussion

The Miami subjects were primarily Cuban-American children of Cuban refugees after the Cuban revolution. Although there are no specific data available on the socioeconomic status of these subjects, they tended to come from more highly educated Hispanic families. These children tested more similar to the monolingual samples than did the Tucson-San Diego or Hayward group. To find bilingual children whose families speak Spanish in the home, it was necessary to go to areas of more recent immigrants. These families tended to be from lower socioeconomic status homes. It must be noted that the majority of the Hayward subjects came from lower socioeconomic status homes.

In summary the results of this study appear to warrant the following conclusions:

1. Bilingual children scored as well or better on the English translation of the Spanish test. The kindergarten children scored significantly higher on the English translation on two of the ten subtests but not on the total score. The second

grade children scored significantly higher on the English translation on four of the ten subtests and on the total score.

2. The differences between the English and Spanish test appear to be primarily on tests in the auditory-vocal channel of communication rather than on tests in the visual motor channel.
3. Testing order was controlled by the design of administering the Spanish test first to one-half the subjects and the English translation first to the other half. It was found that there were practice effects on the second testing for several of the ITPA subtests in both Spanish and English. Statistically, this did not have any effect on the main results.
4. The mean discrepancy across the profile was calculated to determine the numbers of children who could be classified as severely learning disabled according to whether they were tested in Spanish or English. There appears to be little difference in the extent of the mean deviation discrepancies of scores for the kindergarten children regardless of the language used in the test. By second grade the results indicate that only one child showed significant deviations when tested in English and six had significant deviations when tested in Spanish. These deviations were the result of low scores on some auditory-vocal tests. Therefore, testing second grade children in Spanish may erroneously identify children as learning disabled.
5. Comparisons were made between the monolingual children in the standardization sample and samples in Hayward, San Diego-Tucson, and Miami. The Hayward sample was found to score lower than the

monolinguals on most auditory-vocal subtests and similar to the San Diego-Tucson sample. However, the Hayward children scored lower than the Miami sample.

CHAPTER 5

SUMMARY AND CONCLUSIONS

The purpose of this study was to determine whether bilingual children score higher on the Spanish standardized ITPA or its English translation. The significance of this problem has been increased by recent court decisions requiring that bilingual Hispanic children be tested in their primary language.

Procedure

Twenty-six kindergarten and 28 second grade bilingual (Spanish/English) children enrolled in school in a California community were tested on the Spanish standardized Illinois Test of Psycholinguistic Abilities and its English translation. All children were Spanish-American and reported that Spanish was spoken in their home. All were in bilingual programs and identified by teacher judgment as "bilingual." Slightly more than half the kindergarten subjects and second grade subjects were female. Kindergarten subjects were all five and six years old, whereas second grade subjects were seven, eight, and nine years old. None of the subjects had hearing, vision or other physical handicaps. Approximately 80 percent of the mothers of kindergarten subjects and second grade subjects were born in Spanish-speaking countries. The remaining mothers were born in the United States. Eighty-five percent of the fathers of kindergarten and second grade subjects were born in

Spanish-speaking countries with the remaining 15 percent born in the United States. Eighty percent of the subjects were born in California. The remaining 20 percent were born in Mexico. Approximately 28 percent of the subjects were from middle socioeconomic status families and 72 percent from lower socioeconomic status families. About one-half of the kindergarten and one-half of the second grade children had attended preschool. All families of subjects in this study considered themselves as Spanish-American.

To select the subjects for the study, the master computer list maintained by the district was utilized to gain the names of children involved in the bilingual program who were orally fluent in both Spanish and English. The bilingual program assistants (who supervise the bilingual teachers), and who are familiar with the children, teachers, and schools selected those children they felt were "most" bilingual. All of the kindergartners and some of the second graders had been administered the Bilingual Syntax Measure yielding a score in both languages. The San Diego Observation Assessment Instrument had been administered to many of the students. Because no reliable instrument exists for measuring the degree of bilingualism, scores on a bilingual test were not used as part of the criterion for selecting children. The following criteria were used in selecting the sample of bilingual children.

1. All children were in the bilingual program on a voluntary basis.
2. Teacher and program assistant both agreed that the child had developed expressive and receptive use of both Spanish and English.

3. The children and parents reported that Spanish is spoken in the home of the child.

All subjects were tested on the Spanish standardized ITPA and its English translation. The split-half method was used. Half of the children at each grade level were tested first in Spanish and then tested within four days in English. The other half of the subjects were tested in English first and then in Spanish within four days of the English administration. All communications with the subjects were conducted in the language in which the child was tested. All tests were administered by the author, who had participated in the development of the Spanish edition of the ITPA.

Results

The results of this study were:

1. Bilingual children scored as well or better on the English translation of the Spanish ITPA. The kindergarten children scored significantly higher on the English translation on two of the ten subtests. The second grade children scored significantly higher on the English translation on four of the ten subtests and on the total score.
2. The differences between the performance of the children on the English and Spanish tests appeared to be primarily on tests in the auditory-vocal channel rather than on tests in the visual motor channel.
3. Testing order was controlled by the design of administering the Spanish test first to one-half the subjects and the English

translation first to the other half. It was found that there were practice effects on the second testing for several of the ITPA subtests in both Spanish and English. Statistically, this did not have any effect on the main results.

4. The mean deviation score across the ITPA was calculated for each child to determine the numbers of children who would be classified as severely learning disabled on the Spanish test or on the English test. There appeared to be little difference in the deviation scores for the kindergarten children regardless of the language used in the test. However, by second grade, only one child showed significant deviations when tested in Spanish. These deviations were the result of low scores on some auditory-vocal tests. Therefore, testing second grade children in Spanish may erroneously identify children as learning disabled.
5. Comparisons were made between the monolingual children in the standardization sample and samples of bilingual children in Hayward, Tucson-San Diego and Miami. The Hayward sample was found to score lower than the monolinguals on most auditory-vocal subtests and similar to the Tucson-San Diego sample. The Hayward children scored lower than the Miami sample on all subtests of the ITPA.

In summary, this research indicates that bilingual children in this sample scored slightly higher on an English translation of the Spanish ITPA, standardized on and adapted to a Spanish-speaking population.

Discussion and Implications

In reviewing the literature it was found that the question of whether or not bilingual children are penalized on individual verbal and nonverbal intelligence tests when tested in English has not been answered. The research to date has been equivocal. Early research used group tests which were an inadequate means of assessing intelligence and comparing individual children. The majority of researchers using individually administered intelligence tests tested children in English only. These researchers found bilingual children penalized on individually administered verbal tests of intelligence (Seidl, 1937; Darcy, 1946; Roca, 1955). Studies testing children in both languages found different conclusions (Anastasia and Cordova, 1953; Hickey, 1972; Levandowski, 1975). These studies, however, used a translation of a test developed and standardized in English.

Court decisions in recent years have mandated the development of bilingual programs for children coming from homes where English is not the primary language. Lau vs. Nichols (1973) was an important case involving Chinese-American children in San Francisco. In this case the courts not only mandated bilingual education but set itself up as a mediator should enforcement difficulties develop. Other court decisions mandated still other services. As an example, in Guadalupe vs. Tempe Elementary School District (1972), the courts ruled that bilingual children placed in special education had to be tested in their primary language. In Diana vs. State Board of Education (1970) in California, the court ordered the retesting of over 22,000 Mexican-American children who had previously been placed in special education on the basis of

English language intelligence tests. Clearly, the courts have mandated the testing of bilingual children in their primary language.

The results of this investigation do not support the court decision to test children in their native language. When a test standardized on Spanish-speaking children (the Spanish ITPA) was developed and bilingual children were tested with this instrument and the English translation, it was found that kindergarten and second grade children actually performed better in English than in Spanish. In view of the fact that results of this study are contrary to popular opinion and current court rulings, it is important to raise several questions about the study:

1. Is the sample of this study representative of bilingual children? The subjects of this study came from eight elementary schools in Hayward, California. The children were Mexican-American. One must tentatively conclude that the results apply only to the population sample from which the sample for this study was drawn. Efforts were made to select a sample that was truly bilingual, but no comparisons with other communities were made. However, as has been indicated earlier, the Hayward sample was similar to samples of children tested in Tucson, Arizona and San Diego, California. This researcher believes the sample is typical of children classified in schools as bilingual.

2. Is the translation adequate? The Spanish standardized ITPA was translated by the examiner, a bilingual with Puerto-Rican background. Other professional bilinguals helped in the translation. Many of the subtests required little translation. As an example, Verbal Fluency, (which showed a superiority for the English translation at the

kindergarten and second grade levels) requires a brief set of directions given before the timing begins and the child says all the words he can. No other instructions are necessary. The Spanish and English translation of the test are available from the author. The simplicity of the language, aimed at children aged 2-10, reduced problems of translation.

3. Was the data properly analyzed? The data were placed on computer cards that were checked several times for accuracy by a statistician familiar with data analysis, but not otherwise involved in the study. The analysis of variance was done several times to insure accuracy and finally, analyses were undertaken by hand using the original raw data and the same results were obtained.

Further Research

Further research is needed to validate or contradict the evidence presented in this study. Research should be conducted using the Spanish ITPA and an English translation on bilingual children in other parts of the country. It may be necessary to have norms for various regions or populations of the country.

Researchers may want to develop and standardize other tests in Spanish and test bilingual children on those tests and the English translations. This would be particularly beneficial if the instruments tested areas different than the subtests of the ITPA. For example, measures of affective responses may yield different results than measures of cognitive functioning. Some bilinguals dream in Spanish, even when immersed in an English-speaking culture and feel culturally

and emotionally different from Anglos. Further research into affective variables may indicate that bilingual children do better on Spanish tests of personality or social competence which do not rely on cognitive or psycholinguistic processes.

Further research should include use or development of an assessment instrument to determine the degree of bilingualism. It is imperative that a test to measure the degree of bilingualism be developed since no reliable or valid test exists at present. No present tests of bilingualism currently meet reliability or validity standards. If such a test can be developed, it may be possible to determine the degree of competency necessary for a child to profit from instruction in English. The important issue is the need for an adequate examination of linguistic competency of children in English and Spanish to determine the degree of bilingualism. Children should be tested in the language in which they are most competent, but it should be remembered that the degree of bilingualism in young children changes rapidly in early grades.

Refinement of an instrument to measure bilingualism is important but does not solve the problem of what degree of bilingualism in Spanish and English is appropriate for classification as "bilingual." Also, it must be remembered that the degree of bilingualism may change quickly in a young child. The results of this study clearly show that second grade bilingual children were more competent in English than kindergarten children. This change varies from child to child, reflecting total ecological impact.

The issue of the language in which a child should be taught underlies a basic controversy in the field of bilingual education. The results of this study do not address this issue and are limited to the question of assessment of the psycholinguistic abilities measured by the ITPA. They should not be interpreted as being nonsupportive of bilingual programs. Further research is necessary to provide the basis for resolving the issues in the field of bilingual education, especially with the handicapped.

Although the results of this study are restricted to Spanish-English speaking children, it raises questions about assessment of children who speak other languages, e.g., recent immigrants from Vietnam, China, Laos, etc., or children in the French immersion programs in Quebec. Similar studies would be necessary to answer research questions with these populations of bilingual children.

In conclusion, the results of this study have raised questions about the scientific validity of recent court decisions requiring that bilingual children be tested in their primary language.

*Anyone interested in the English translation of the Spanish edition of the ITPA may write to the author.

REFERENCES

Altus, G. J. WISC patterns of a selected sample of bilingual school children. Journal of Genetic Psychology, 1953, 83, 241-248.

Anastasi, A. and Cordova, F. A. Some effects of bilingualism upon the intelligence test performance of Puerto Rican children in New York City. Journal of Educational Psychology, 1953, 44, 1.

Anastasi, A. and de Jesus, C. Language development and nonverbal IQ of Puerto Rican preschool children. Journal of Abnormal and Social Psychology, 1953, 48, 357-366.

Arizona Revised Statutes, Annotated Sec. 15-202, 15-10 7-1099 (Supp. August 1973).

Arsenian, S. Bilingualism in the post-war world. Psychological Bulletin, 1945, 42, 65-86.

Aspira of New York, Inc. v. New York Board of Education, 72 Civ. 4002 (S.D. N.Y., Filed September 20, 1972).

Barbe, E. M. A study of the comparative intelligence of the children in certain bilingual and monolingual schools in South Wales. British Journal of Educational Psychology, 1933, 3, 237-259.

Bilingual Educational Act--Title VII, Elementary and Secondary Education Act of 1965, 20 U.S.C.

Bilingual Educational Act of 1968, 20 U.S.C., Sec. 880b-2(c) (1970).

Brown v. Board of Education, 347 U.S. 483 (1954).

Christiansen, T. and Livermore, G. A. A comparison of Anglo American and Spanish-American children on the WISC. Journal of Social Psychology, 1970, 81, 9-14.

Civil Rights Act of 1964, 601, 42 U.S.C., Sec. 2000d (1970).

Civil Rights Digest, 12, (December 1971).

Colvin, S. S. and Allen, R. D. Mental tests of linguistic ability. Journal of Educational Psychology, 1923, 14, 1-20.

Darcy, N. T. The effect of bilingualism upon the measurement of the intelligence of children of preschool age. Journal of Educational Psychology, 1946, 37, 21-44.

Diana v. State Board of Education, Clearing House Review, Volume 3, #10, February 1970.

Federal Register, Section 11595 (1970).

Guadalupe Organization Inc. v. Tempe Elementary School District, No. CIV 71-435 Phx. (D. Arizona, January 24, 1972).

Hickey, T. Bilingualism and the measurement of intelligence and verbal learning ability. Exceptional Children, 1972, 39, 24-28.

Hill, H. S. Correlation between IQs of bilinguals at different ages on different intelligence tests. School and Society, 1936, 44, 89-90.

Hirsch, N. D. A study of natio-racial mental differences. Genetic Psychological Monographs, 1926, 1, 231-407.

Jamieson, I. and Sandiford, P. The mental capacity of Southern Ontario Indians. Journal of Educational Psychology, 1928, 19, 313-328.

Keys v. School District, 445 F. 2d 990 (CA 10, 1973).

Kirk, S. A. and Elkins, J. Ethnic differences and learning disabilities in head start children. Tucson, Arizona: Learning Disabilities Leadership Training Institute, Department of Special Education, The University of Arizona, June 1974.

Kirk, S. A., McCarthy, J. J. and Kirk, W. D. Examiner's Manual. Illinois Test of Psycholinguistic Abilities (Revised Edition). Urbana, Illinois: University of Illinois Press, 1968.

Kittell, J. E. Intelligence test performance of children from bilingual environments. Elementary School Journal, 1963, 64, 76-83.

Kobrick, A. A model act provided for transitional bilingual educational programs in public schools. Harvard Journal of Legislation, 1972, 260-264.

Lau et al. v. Nichols et al., 414 U.S., p. 563 (1973).

Leopold, W. F. Speech Development of a Bilingual Child. Evanston, Illinois: Northwestern University Press, 1939.

Lester, O. F. Performance tests of foreign children. Journal of Educational Psychology, 1923, 20, 303-309.

Levandowski, B. The difference in intelligence scores of bilingual students on an English version of the intelligence test as compared to a Spanish version of the test. Illinois School Research, Vol. 11, No. 3, Spring 1975.

Levinson, B. M. A comparison of the performance of bilingual and monolingual native-born Jewish preschool children of traditional parentage on four intelligence tests. Journal of Clinical Psychology, 1959, 15, 74-76.

McCarthy, J. J. and Kirk, S. A. The Illinois Test of Psycholinguistic Abilities, Experimental Edition, Examiner's Manual. Urbana, Illinois: University of Illinois Press, 1961.

Mead, M. Group intelligence and linguistic disability among Italian children. School and Society, 1927, 25, 465-468.

Mitchell, A. J. The effects of bilingualism in the measurement of intelligence. Elementary School Journal, 1937, 38, 29-37.

Ortiz, A. Assessing the academic and social performance of Mexican-American children. Presentation at International Conference for the Association for Children with Learning Disabilities. San Francisco: March 3, 1979.

Osgood, C. E. Motivational dynamics of language behavior. In M. Jones (Ed.) Nebraska Symposium of Motivation. Lincoln, Nebraska: University of Nebraska Press, 1957, 348-424.

Palmer, M. and Gaffney, P. D. Effects of administration of the WISC in Spanish and English and relationship of social class to performance. Psychology in the Schools, 1972, 9, 1.

Paraskevopoulos, J. and Kirk, S. A. The Development and Psychometric Characteristics of the Revised Illinois Test of Psycholinguistic Abilities. Urbana, Illinois: University of Illinois Press, 1968.

Pintner, R. Comparison of American and foreign children on intelligence tests. Journal of Educational Psychology, 1923, 14, 292-295.

Pintner, R. The influence of language background on intelligence tests. Journal of Social Psychology, 1932, 3, 235-240.

Pintner, R. and Keller, R. Intelligence of foreign children. Journal of Educational Psychology, 1922, 13, 214-222.

Rigg, M. Some further data on language handicap. Journal of Educational Psychology, 1928, 19, 252-256.

Roca, P. Problems of adapting intelligence scales from one culture to another. High School Journal, 1955, 18, 124-131.

Seidl, J. C. The effect of bilingualism on the measurement of intelligence. Unpublished doctoral dissertation. New York: Fordham University, 1937.

Serna v. Portales Municipal Schools, 351 F Supp. 1279 (D. N.M., 1973).

Spoerl, P. T. The academic and verbal adjustment of college-age bilingual students. Journal of Genetic Psychology, 1943, 64, 139-157.

Stark, W. A. The effect of bilingualism on general intelligence: an investigation carried out in certain Dublin primary schools. British Journal of Educational Psychology, 1940, 10, 78-79.

Swanson, M. M. Bilingual Education: The National Perspective. Skokie, Illinois: National Textbook Co., 1974.

U.S. v. Texas, 342 F. Supp. 24 (E.D. Tex. 1971); aff'd, 466 F. 2d 518 (5th Cir. 1972).

Wepman, J. M. Auditory Discrimination Test. Chicago: Language Research Associates, 1958.

Yoshioka, J. G. A study of bilingualism. Journal of Genetic Psychology, 1929, 36, 473-479.

Zirkel, P. A. Spanish-speaking students and standardized tests. The Urban Review, 1972, 5, 33-37.