THE RELATIONSHIP OF SEX AND AGE AT ENTRANCE TO SCHOOL TO SECOND GRADE ACHIEVEMENT

DISSERTATION

Presented to the Graduate Council of the North Texas State University in Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF EDUCATION

By

Sharon Reynolds Jernigan, B.A., M.Ed.
Denton, Texas
December, 1986
This investigation compared achievement of boys and girls in second grade who were seven years old in June, July, and August of 1983 to the boys and girls in second grade who were eight years old in September, October, and November of 1983. The students were tested using the Iowa Test of Basic Skills using the following areas: reading, total math, and composite scores. The study also looked at the correlation of sex and age of students who had been retained in first grade. A comparison of teacher grades to standardized test scores and ability grouping was also presented.

One way analysis of variance was applied to the test results. A chi square test of independence was conducted on students retained in the first grade to determine if interaction between sex and age was indicated.

Older children scored higher in all three areas measured, while girls scored higher in reading. This may seem contradictory, but is not. Age was significant beyond the .05 level, while sex was significant beyond the .001 level. This difference is explained by the extremes in means for
younger boys and older girls. Since first grade curriculum emphasizes reading, this gave girls a definite advantage over boys. Boys, however, scored significantly higher in math. The results indicate a need for restructuring curriculum to meet the needs of boys and girls.

Younger boys in second grade scored the lowest in all areas tested, except math. These scores would have been even more significant if the boys who were not promoted to second grade could have been included in the second grade testing.

These findings indicate that total developmental age is the most important factor when considering admission for school. A closer look should be taken at the requirements for school entrance. The factors that must be considered are chronological age, mental age, physical maturity, emotional and social maturity, behavior age, and sex.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td></td>
</tr>
<tr>
<td>Hypotheses and Research Questions</td>
<td></td>
</tr>
<tr>
<td>Significance of the Study</td>
<td></td>
</tr>
<tr>
<td>Definition of Terms</td>
<td></td>
</tr>
<tr>
<td>Basic Assumptions</td>
<td></td>
</tr>
<tr>
<td>Selection of the District</td>
<td></td>
</tr>
<tr>
<td>Procedures for Collection of Data</td>
<td></td>
</tr>
<tr>
<td>Procedures for Analysis of Data</td>
<td></td>
</tr>
<tr>
<td>II. REVIEW OF RELATED LITERATURE</td>
<td>10</td>
</tr>
<tr>
<td>History</td>
<td></td>
</tr>
<tr>
<td>Chronological Age</td>
<td></td>
</tr>
<tr>
<td>Developmental Age</td>
<td></td>
</tr>
<tr>
<td>Sex and Age</td>
<td></td>
</tr>
<tr>
<td>Effects of Failure</td>
<td></td>
</tr>
<tr>
<td>Current Practices</td>
<td></td>
</tr>
<tr>
<td>Teacher Grades</td>
<td></td>
</tr>
<tr>
<td>III. PROCEDURES FOR COLLECTING AND TREATING THE DATA</td>
<td>26</td>
</tr>
<tr>
<td>Description of the Students</td>
<td></td>
</tr>
<tr>
<td>Selection of the Students</td>
<td></td>
</tr>
<tr>
<td>Instrumentation</td>
<td></td>
</tr>
<tr>
<td>Procedures for Collection of Data</td>
<td></td>
</tr>
<tr>
<td>Procedures for Analysis of Data</td>
<td></td>
</tr>
<tr>
<td>IV. PRESENTATION AND ANALYSIS OF THE DATA</td>
<td>33</td>
</tr>
<tr>
<td>Hypothesis One</td>
<td></td>
</tr>
<tr>
<td>Hypothesis Two</td>
<td></td>
</tr>
<tr>
<td>Hypothesis Three</td>
<td></td>
</tr>
<tr>
<td>Hypothesis Four</td>
<td></td>
</tr>
<tr>
<td>Hypothesis Five</td>
<td></td>
</tr>
<tr>
<td>Hypothesis Six</td>
<td></td>
</tr>
<tr>
<td>Hypothesis Seven</td>
<td></td>
</tr>
</tbody>
</table>
## TABLE OF CONTENTS--Continued

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypothesis Eight</td>
<td></td>
</tr>
<tr>
<td>Hypothesis Nine</td>
<td></td>
</tr>
<tr>
<td>Summary of Findings</td>
<td></td>
</tr>
<tr>
<td>V. FINDINGS, CONCLUSIONS AND RECOMMENDATIONS . . . . . . . . . . . . .</td>
<td>63</td>
</tr>
<tr>
<td>Findings</td>
<td></td>
</tr>
<tr>
<td>Discussion of Findings</td>
<td></td>
</tr>
<tr>
<td>Conclusions</td>
<td></td>
</tr>
<tr>
<td>Recommendations</td>
<td></td>
</tr>
<tr>
<td>APPENDIX</td>
<td>77</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>78</td>
</tr>
<tr>
<td>Table</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>I. Sum of Squares, Degrees of Freedom, Mean Squares, F, and Significance of F for Reading Scores for Older and Younger Students on the Iowa Test of Basic Skills</td>
<td>35</td>
</tr>
<tr>
<td>II. Sum of Squares, Degrees of Freedom, Mean Squares, F, and Significance of F for the Total Math Scores for Older and Younger Students on the Iowa Test of Basic Skills</td>
<td>36</td>
</tr>
<tr>
<td>III. Sum of Squares, Degrees of Freedom, Mean Squares, F, and Significance of F for the Composite Scores for Older and Younger Students on the Iowa Test of Basic Skills</td>
<td>37</td>
</tr>
<tr>
<td>IV. Sum of Squares, Degrees of Freedom, Mean Squares, F, and Significance of F for Reading Scores for Boys and Girls on the Iowa Test of Basic Skills</td>
<td>39</td>
</tr>
<tr>
<td>V. Sum of Squares, Degrees of Freedom, Mean Squares, F, and Significance of F for Total Math Scores for Boys and Girls on the Iowa Test of Basic Skills</td>
<td>40</td>
</tr>
<tr>
<td>VI. Sum of Squares, Degrees of Freedom, Mean Squares, F, and Significance of F for Composite Scores for Boys and Girls on the Iowa Test of Basic Skills</td>
<td>41</td>
</tr>
<tr>
<td>VII. Sum of Squares, Degrees of Freedom, Mean Squares, F, and Significance of F for Reading Subtest for Boys and Girls on the Iowa Test of Basic Skills</td>
<td>43</td>
</tr>
<tr>
<td>VIII. Means for Younger Boys and Girls and Older Boys and Girls on the Reading Subtest of the Iowa Test of Basic Skills</td>
<td>44</td>
</tr>
</tbody>
</table>
LIST OF TABLES—Continued

Table | Description | Page
--- | --- | ---
IX. | Sum of Squares, Degrees of Freedom, Mean Squares, F, and Significance of F for the Total Math Scores for Boys and Girls on the Iowa Test of Basic Skills | 45
X. | Means for Younger Boys and Girls and Older Boys and Girls on the Total Math Scores of the Iowa Test of Basic Skills | 46
XI. | Sum of Squares, Degrees of Freedom, Mean Squares, F, and Significance of F for the Composite Scores for Boys and Girls on the Iowa Test of Basic Skills | 47
XII. | Means for Younger Boys and Girls and Older Boys and Girls on the Composite Scores of the Iowa Test of Basic Skills | 48
XIII. | Number of Students Retained in First Grade | 55
XIV. | Year End Grades and Ability Groups for First Grade Students From One Elementary School | 58
CHAPTER I

INTRODUCTION

Even though research has continued to show that early admission to school results in higher failure rate, admission to school has continued to be based on the arbitrarily chosen chronological age of six for first grade throughout the history of American education (Ames & Ilg, 1979; Barrett, 1956; Durrell, 1958; Dykstra, 1966; Elkind, 1986). The minutes of the Dartmouth, Massachusetts, School Committee in 1843 recorded that children under four years of age were being admitted. The voters of Dartmouth began their drive for a minimum chronological age for school entrance because of a high incidence of failures (Braga, 1971; Comisky, 1957; Hedges, 1977). Finally, in 1909 a chronological age of six years was set in Dartmouth as the minimum standard. It is not clear as to why age six became the standard in Dartmouth as well as other school systems after the advent of compulsory attendance laws in 1852 (Braga, 1971; Hedges, 1977).

At the turn of the century, educators began to challenge the practice of a fixed age for school admission. For example, Dewey and Kilpatrick believed the high incidence of failure at all grade levels was due to
children beginning school before they were developmentally ready. Evidence has continued to accumulate supporting readiness as a major factor in school success (Rowland & Nelson, 1959).

By the 1940's Gesell and others were beginning to conduct research on factors facilitating success in first grade (Gessell, 1940). The common thread throughout the research was that large numbers of children were failing in first grade. One source of failure was related to immaturity (Hedges, 1977).

Formal instruction puts excessive demands on young children. Those who are not ready for the demands of school structure experience unnecessary pressure (Doll, 1966; Elkind, 1984). Pediatricians from different areas of the United States have reported an increase in stress-related symptoms, such as peptic ulcers, in young children (Brazelton, 1985).

Although there have been several studies conducted since the 1950's that compare achievement of younger children to older children in the same grade level and research that indicates girls achieve higher than boys, this researcher was unable to locate research that compares the interrelationship of both age and sex to achievement. This study is different in that it does compare the relationship of sex and age at entrance to school to three different achievement test scores.
Purpose of the Study

The purpose of this study was to determine the relationship of sex and age at entrance to school on second grade achievement as measured by the Iowa Test of Basic Skills (Form 7, Level 7) in the following areas: reading, total math, and the composite of all subtests.

Hypotheses and Research Questions

1. Second grade students whose birth dates were between September and November in 1975 (older students) will score significantly higher on reading scores, total math scores, and composite scores of the Iowa Test of Basic Skills (Form 7, Level 7) than second grade students whose birth dates were between June and August in 1976 (younger students).

2. Second grade girls will score significantly higher on reading scores, total math scores, and composite scores of the Iowa Test of Basic Skills (Form 7, Level 7) than will second grade boys.

3. Second grade girls whose birth dates were between September and November in 1975 (older girls) will score significantly higher on reading scores, total math scores, and composite scores of the Iowa Test of Basic Skills (Form 7, Level 7) than second grade boys who are the same age (older boys).

4. Second grade girls whose birth dates were between June and August in 1976 (younger girls) will score
significantly higher on reading scores, total math scores, and composite scores of the Iowa Test of Basic Skills (Form 7, Level 7) than second grade boys who are the same age (younger boys).

5. Second grade girls whose birth dates were between September and November in 1975 (older girls) will score significantly higher on reading scores, total math scores, and composite scores of the Iowa Test of Basic Skills (Form 7, Level 7) than second grade girls whose birth dates were between June and August in 1976 (younger girls).

6. Second grade boys whose birth dates were between September and November in 1975 (older boys) will score significantly higher on reading scores, total math scores, and composite scores of the Iowa Test of Basic Skills (Form 7, Level 7) than second grade boys whose birth dates were between June and August in 1976 (younger boys).

7. Second grade girls whose birth dates were between September and November (older girls) will score significantly higher on reading scores, total math scores, and composite scores of the Iowa Test of Basic Skills (Form 7, Level 7) than second grade boys whose birth dates were between June and August in 1976 (younger boys).

8. Second grade boys whose birth dates were between September and November in 1975 (older boys) will not score significantly higher on reading scores, total math scores,
and composite scores of the Iowa Test of Basic Skills (Form 7, Level 7) than will second grade girls whose birth dates were between June and August in 1976 (younger girls).

9. Of those children retained in first grade, there will be more younger boys (birth dates between June and August, 1976) than any other group.

10. The following research questions will be discussed:
   a. Are teacher grades valid in evaluating student achievement at second grade level?
   b. Do teacher grades correlate with the scores of the Iowa Test of Basic Skills?

**Significance of the Study**

Historically, an arbitrary chronological age has been the major requirement for school admission. Even though the literature shows that there is a difference in the development of boys and girls, most public schools have never recognized this difference as a determining factor for allowing a child to enter school. If the results of this study are significant, then public schools will need to provide alternatives to using chronological age as the criterion for admission to school.

**Definition of Terms**

1. Older boys--Boys whose birth dates were between September and November in 1975.
2. **Older girls**—Girls whose birth dates were between September and November in 1975.

3. **Younger boys**—Boys whose birth dates were between June and August in 1976.

4. **Younger girls**—Girls whose birth dates were between June and August in 1976.

5. **ITBS**—Iowa Test of Basic Skills, Form 7, Level 7.


7. **Reading Score**—A test of reading comprehension on the ITBS.

8. **Total Math Score**—A score combining Mathematics Concepts, Mathematics Problems, and Mathematics Computation on the ITBS.

**Basic Assumptions**

The major assumptions underlying this study were that the Iowa Test of Basic Skills (Form 7, Level 7) was given to second grade students in a controlled setting and that the test was a valid reflection of achievement.

**Selection of the District**

The district in which this study was performed was selected because of an agreement between the district and North Texas State University allowing research within the
district. Second grade test scores were taken from every elementary school in the district.

Procedures for Collection of Data

Sample

The final list of students in the study was taken from the entire first and second grade population of elementary schools in Garland, Texas. A computer list that contained sex and birth dates of all first and second grade students was generated. Second grade students chosen for the study had birth dates in June, July, and August of 1976 and September, October, and November of 1975. These students were then divided into four groups by age and sex. First grade students in the study were those who by chronological age should have been in second grade. This was determined by birth dates. There was a total of 803 second grade students and 175 students retained in first grade in the study.

Instrumentation

The school district administered the Iowa Test of Basic Skills (Form 7, Level 7) to all second grade students in September of 1983. It provided student scores for this study.
Procedures for Analysis of Data

One way analysis of variance was used to determine differences among group means. Three tests of significance were run. The first used the scores from the reading subtest, the second used the combined scores from the math subtests, and the third used the composite score of all subtests. Multiple comparisons using Scheffe's method were conducted, where appropriate, to determine which group means differed.

Chi square was the statistical procedure used in the analysis of the students who were retained in first grade. A chi square test of independence was conducted to determine if interaction was indicated.


CHAPTER II

REVIEW OF RELATED LITERATURE

History

Historically school admission has been based on chronological age. After the advent of compulsory attendance laws, many school districts set the chronological age of six as the minimum standard (Braga, 1971; Hedges, 1977).

Evidence accumulated since that time has challenged the practice of a fixed age for school admission (Rowland & Nelson, 1959). A common thread throughout the research was that failure in first grade was related to immaturity (Hedges, 1977). A study conducted in suburban Long Island in 1960-1961 compared achievement of pupils admitted to first grade at an early age to those of pupils admitted at a later age. The comparison was made at the end of fourth and fifth grades. Older fourth grade pupils were superior at all levels except arithmetic fundamentals. The same was true at the end of fifth grade (Halliwell & Stein, 1964).

The history of preschool screening for developmental disabilities has its roots in the fields of medicine and psychology. One of the earliest recorders of infant development was Charles Darwin, who produced one of the first biographical reports on this subject in the late nineteenth
century (Darwin, 1877). Another pioneer in this area was Arnold Gesell who created and ultimately published the Gesell Developmental Schedules (Gesell, 1940; Gesell & Amatruda, 1941).

**Chronological Age**

There is little doubt that certain children are ready for printed words long before the age of six. On the other hand, there are those children whose age and intelligence entitle them to entrance to school, but who do not respond to academic instruction. For both of these groups of children chronological age alone is a poor predictor. It is possible that chronological age should be viewed in relation to sex. Among the "superior immatures" (children who have a high mental age but low social skills), the majority are boys (Jansky & DeHirsch, 1972).

Durrell (1958), Barrett (1956), and Dykstra (1966) also found chronological age wanting as a predictor. Their results indicate that a study of the interaction between age and sex might be enlightening because of the frequent failure of young boys. Singer (1968) suggests that different tests or perhaps different scoring systems may be needed for boys and girls.

No child fits all the norms for his or her age. Each will have some behavior typical of younger children and some typical of older children (Levenson, 1977). Young children
do not learn in the same ways as older children and adults (Elkind, 1986). Because of this difference, chronological age is not a meaningful predictor of school success (Ames & Ilg, 1979; Hedges, 1977). There is no single criterion for the determination of school entrance that adequately predicts school success. The factors that must be considered include chronological age, mental age, physical maturity, emotional and social maturity, behavior age, and sex (Braga, 1971; Brenner, 1957). Most educators know about this discrepancy in behavior level and rate of development of the two sexes, but almost nobody does anything practical about it (Ames, 1972; Hedges, 1977).

R. Vance Hall (1963) conducted a study of the relationship of school entrance age of boys and girls to school achievement. In this study he found that 801 out of about 12,800 elementary pupils had been retained at one grade level. Almost three-fourths of these were boys. Of those who were retained, about 78 percent of the boys and 80 percent of the girls were underage when they started the first grade.

Also in 1963, a study was conducted to determine the effect of chronological age at the time of entering school on later achievement. The information was collected in the Sioux Falls Public School System. All children in the study met requirements for entrance to school. The younger one-fourth of the children had a significantly lower mean
composite score of the Iowa Test of Basic Skills than the remainder of the class (Dickinson & Larson, 1963).

A study conducted in Montgomery County (Maryland) public schools involving an early admission program indicated that very few five year olds are ready for first grade. The evaluation also indicated the need for multiple criteria, such as mental age and social development, for identifying children for placement (Hebbeler, 1983).

According to Hedges (1977), children exactly six years old will vary as much as 2.3 years in mental age. Studies done and replicated several times in Winnetka public schools conclude that if the teaching of reading were postponed until children had attained a mental age of six and one-half years, the chances of failure would be decreased (Hedges, 1977). The advantages of postponing early entrance to first grade programs as they are presently conducted are very real (Halliwell, 1966). Children who are younger at school entrance often have academic problems that last throughout their school careers (Uphoff & Gilmore, 1985).

Moore and Moore (1975) also found scant evidence in research for lasting benefits for most children from early schooling. The scientific evidence impressively favors later school admission. Formal instruction puts excessive demands on young children. Those who are not ready for the demands of school structure experience unnecessary pressure (Doll, 1966; Elkind, 1984). Pediatricians from different
areas of the United States have reported an increase in stress-related symptoms, such as peptic ulcers, in young children (Brazelton, 1985).

Developmental Age

Studies conducted at the Gesell Institute indicate that maturation is a slow process (Gesell & Ilg, 1946). While the general course of development is similar for boys and girls, the rate of development for boys appears slower than for girls (Ames, 1972; Dechant, 1968; DiPasquale, 1980; Gesell & Ilg, 1946; Gordon, 1966; Packard, 1968; Pauley, 1952; Pullen, 1972). In the average first grade, there is a spread of four years in the readiness of all students to learn (Goodlad & Anderson, 1969).

The Gesell Developmental Test (The Gesell Institute, 1965), a means for determining behavior age, has been used extensively in research. In an initial study of 100 kindergarten children the investigators' judgments as to readiness were in close agreement with teacher judgments, and also were highly related to academic section assignments six years later. Developmental assessment should be a continual process (Elkind, 1976). In his review of the Gesell Developmental Test, Borstelmann (1972) believed this test to be a valuable resource for screening procedures for school entry and initial class assignments.
The majority of children referred to the Gesell Institute for clinical service were doing poorly in school, were overplaced, and were underendowed for the schoolwork being demanded of them. Approximately five times as many boys as girls seem to find themselves in severe enough trouble in school to be taken by their parents for clinic help (Ames, 1972). Young children are more likely to be referred for academic problems than older children (DiPasquale, 1980).

Research conducted at the Gesell Institute reveals that a high I.Q. or above average reading ability, unless it is supported by a total performance which is fully at age (behavior level), is no more effective than chronological age in guaranteeing success in school. The level of intelligence and the level of maturity are two separate variables. It is possible to be superior but immature (Ames, 1966). A child's developmental maturity is the key predictor of success in school (Freisen, 1984).

Sex and Age

It is contrary to the laws of nature that children should advance lock step through the grades (Beck, 1973). Girls seem to have an inherited language advantage or reach maturity in this respect about a year and a half earlier than boys. Girls as a group achieve better than boys in reading (Dechant, 1968).
Pauley (1959) conducted research in Tulsa, Oklahoma, involving 29,992 children attending grades two through eight in 1956-1957. From this study he perceived the differences in achievement between boys and girls and between older children and younger children at the same level.

A study was conducted by Good (1981) using a random sample taken from first grade reading scores on the Iowa Test of Basic Skills. The older students scored significantly higher than the younger students beyond the .01 level. Girls as a group scored significantly higher than boys as a group beyond the .01 level.

Many studies relating age to achievement fail to analyze data separately for males and females. This hinders comparability among investigations (Laven, 1965).

Effects of Failure

School is a complex society in itself. Many different demands are made on children. Success in school means that a child can cope physically, socially, and emotionally and without undue stress (Levenson, 1977). Moore and Moore (1975) believed that states that required school entrance at age five or six for all children are not being fair to them, especially to boys who at the same age are not as developmentally mature as girls.
Elkind (1982) believed that pressure by parents for children to read at an early age reflected parental needs, not the children's needs or inclinations. Most children do as well as they can, and would do better if they could (Ames & Gillespie, 1970). Studies suggest that children confronted with the task of learning to read before they have the requisite mental abilities can develop long-term learning difficulties. They cannot handle the day-to-day learning tasks and fall behind. They become frustrated and apathetic toward school in general. According to Erikson (1963), failure at this stage of development would serve to promote feelings of inferiority. Most often the child defined as having specific learning disabilities has average or above average intelligence, is usually male, and often has well-educated parents (Cline & Ishee, 1972).

Both research and government efforts to raise the intelligence quotients of bright or not so bright children, or to improve the reading ability of the non-reader, have been disappointing ineffective. This indicates that any positive effects of pushing are seldom lasting (Ames & Chase, 1974).

Current Practices

The majority of schools continue to enter children on the basis of their chronological age. This age requirement varies from state to state, and seems often to depend more
on legislative actions than on research conducted concerning the effects of the age requirement on the children (Ames, 1966; Hedges, 1977).

There are those who disagree with the idea that more time for maturation helps achievement in school. New York State's Commissioner of Education recently proposed that kindergarteners begin at age four and that the public school curriculum be completed at the end of the eleventh grade.

Almost a decade ago the California State Department of Education in its Report of the Task Force on Early Childhood Education proposed that California establish a broadly based educational program that extended one year below the system in existence. White (1972) disagreed with the report that all children four years of age and older receive formal schooling. He said that this recommendation was based on a seriously erroneous interpretation of research evidence and was directly contradicted by that evidence.

In reports to the Texas Select Committee on Public Education chaired by H. Ross Perot, Mortimer Adler (Fish, 1983) and John Goodlad (Stutz, 1983) have both suggested that students enter school at four years of age and graduate by age sixteen. Goodlad suggested that the change should be implemented on a pilot basis and gradually expanded to all children starting school.
Teacher Grades

Over the years experienced teachers have questioned the need for formal readiness tests. In 1947, a study of teachers' ability to subjectively predict first grade success in reading was undertaken by Kottmeyer. The results of this investigation indicated that in the specific area of reading teacher judgment was as valid as results obtained by standardized tests and teachers with more than ten years experience predicted reading success with greater accuracy than those with less experience.

A study conducted by Morgan (1960) suggests that with in-service training kindergarten teachers can predict with reasonable competence the achievement placement of first grade children.

Another study was conducted by Kermoian (1962) to determine the validity of teacher judgment of the readiness status of children entering first grade. The conclusions of this study indicate that the classroom teacher's appraisal of pupil readiness for first grade work correlates highly and significantly with the results of The Metropolitan Readiness Tests.

Other studies have shown that standardized tests place students higher than teacher judgments and higher than their IRI (informal reading inventory) instructional reading levels (Hartman, 1975; Kilgallon, 1942; Liebert, 1965; McCracken, 1963; Oliver & Arnold, 1978).
The research studies mentioned indicate that teachers' judgment is at least as accurate as standardized reading test scores, and perhaps more accurate depending upon the teacher's training and experience. More accurate placement can be determined by coupling standardized test scores, IRI results, and teacher judgment and making adjustments where necessary.
CHAPTER BIBLIOGRAPHY


CHAPTER III

PROCEDURES FOR COLLECTING AND TREATING THE DATA

This chapter presents a description of the subjects, selection of the students, the instrument employed, and the procedures used in analyzing the data.

Description of the Students

The school district selected for this study was located northeast of Dallas, Texas, and had approximately 30,000 students. The socio-economic status ranged from upper lower class to upper middle class. The ethnic composition of the district was 82 percent Anglo, 8 percent Black, 8 percent Spanish surname, and 2 percent Oriental. There were 31 elementary schools in this district.

Selection of the Students

The population for the study was taken from students who entered second grade in the fall of 1983 and those who were retained in first grade for the 1983-1984 school year. There were the following four categories by age and sex for both the first and second grade students in the study:

1. Older boys (199)--birth dates between September and November in 1975.
2. Older girls (199)—birth dates between September and November in 1975.


Instrumentation

The Iowa Test of Basic Skills (Form 7, Level 7) was used to measure achievement. Reliability estimates for the ITBS are acceptably high for Forms 5 and 6 (Harris, 1978). Form 7 is so current that it has not been reviewed.

The publishers of the ITBS report that content specifications are based upon forty years of continuous research in curriculum measurement procedures and interpretation and use of test results. Internal consistency reliability coefficients range from .70 to .95 for the thirteen individual subtest scores. Composite reliability is .97 for first grade and .98 for second grade.

A weakness of the instrument is that a heavy load is placed on the ability to read and understand terms even in some subtests. Depending on what and how a child has been taught in reading, his performance on a subtest could be affected. Harris (1978) and Pyrczak (1978) do agree that in spite of the weaknesses, the ITBS is a satisfactory instrument for obtaining information in the basic skill areas.
The approach used by the Riverside Publishing Company to insure test fairness consists of five stages. These are (1) applications of careful professional judgments at the item writing stage, (2) systematic reviews by representative panels of test users at the assembly stage, (3) statistical analyses based on a variety of bias models, (4) comparisons among the results of stages two and three, and (5) special follow-up studies designed to seek answers raised by the comparisons at stage four. Since any one author's knowledge is necessarily limited, it is common practice to have each test question reviewed by several individuals other than the original item writer before the item is placed in the pool of items to be pretested. To supplement the informal efforts at the item writing and review stage, provisions were made to have all of the items in the recently developed forms of the ITBS reviewed by panels of educators working directly in the kinds of settings in which the tests would be used and representing both sexes and a variety of racial backgrounds.

Wesley and Wesley (1977) state that there is no difference between the mean IQ score for boys and girls. Test designers can manipulate the presence or absence of sex differences. In order to address this kind of situation, a study was conducted in Iowa regarding the ITBS. Data for the study was obtained from forty-eight school systems in Iowa, randomly selected from those who tested third, sixth,
and eighth grade students during the 1975-1976 school year. From each of forty-eight school systems, ten boys and ten girls were randomly selected from each grade level. Vocabulary, Reading Comprehension, Mathematics Concepts, and Mathematics Problem Solving were the subtests used for calculations.

The results of this study suggest that these subtests of ITBS do not function differentially for male and female pupils at the three grade levels tested. The results provide little support for contention that test content as it relates to sex role stereotyping or the frequency of male or female nouns or pronouns, affects performance on mathematics test items. Item location or skill classification did not appear to be a factor in differential item performance for boys and girls. Taken as a whole, therefore, little evidence for sex bias was found in the ITBS (Plake, Hoover, & Loyd, 1980).

The ITBS was standardized jointly with the Cognitive Abilities Test. Approximately 8,000 pupils per grade were used in estimating the spring norms. Fall norms were estimated on representative subsamples of approximately 3,600 in second grade and 2,800 in third grade.

Through continuous review of test materials at all stages of preparation and production by educators of both sexes, an attempt was made to eliminate sex-role
stereotyping, to represent equally the achievements of men and women, to include an approximately equal number of male and female proper nouns, pronouns, and other referents, and to use universal or neutral language to avoid sex-role identification in inappropriate situations (Hieronymus, 1980).

Procedures for Collection of Data

The final list of students in the study was taken from the entire first and second grade population of elementary schools in Garland, Texas. A computer list that contained sex and birth dates of all first and second grade students was generated. Second grade students chosen for the study had birth dates in June, July, and August of 1976 and September, October, and November of 1975. These students were then divided into four sections by age and sex. Grade equivalents for reading, total math, and composite scores on the ITBS were collected for comparison among the groups.

First grade students in the study were those who had not been promoted to second grade. This was determined by birth dates. These students were also divided into four groups and frequencies for sex and age were determined.

There was a total of 803 second grade students and 175 first grade students in the study. A comparison of teacher grades to standardized test scores and ability grouping was also presented.
Procedures for Analysis of Data

One way analysis of variance was the statistical procedure used to test the hypotheses of this study. An analysis was made for the reading scores, total math scores, and composite scores of the Iowa Test of Basic Skills (Form 7, Level 7). Multiple comparisons using Scheffe's method were conducted, where appropriate, to determine which group means differed.

Chi square was the statistical procedure used in the analysis of those students who were retained in first grade. A test of independence was conducted to determine if interaction was indicated.


CHAPTER IV

PRESENTATION AND ANALYSIS OF THE DATA

The relationship of sex and age at school entrance to second grade achievement was analyzed. An analysis was made for the reading scores, total math scores, and composite scores of the Iowa Test of Basic Skills (Form 7, Level 7). The dependent variable was the achievement test score and the independent variables were sex and age.

One way analysis of variance was conducted for hypotheses one through eight as stated in Chapter I. The 5 percent level of significance was used to determine rejection. The Scheffe multiple comparison test was calculated to determine differences in group means. A chi square test of independence was conducted for hypothesis nine.

Hypothesis One

Hypothesis one stated that second grade students whose birth dates were between September and November in 1975 (older students) would score significantly higher on reading scores, total math scores, and composite scores of the Iowa Test of Basic Skills (Form 7, Level 7) than would the second grade students whose birth dates were between June and August in 1976 (younger students). For purposes of analysis and
reporting, the hypothesis was divided into three sub-hypotheses. Hypothesis 1A related age to reading scores. As presented in Table I, the one way analysis of variance indicated an F of 4.20 which was significant beyond the .05 level. Older students scored significantly higher than younger students in reading. Therefore, hypothesis 1A was retained.

Hypothesis 1B related age to total math scores. As shown in Table II, the one way analysis of variance indicated an F of 28.64 which was significant beyond the .001 level. Older students scored significantly higher than younger students in math. Therefore, hypothesis 1B was retained.

Hypothesis 1C related age to composite scores. As shown in Table III, the one way analysis of variance indicated an F of 6.53, which was significant beyond the .001 level. Older students scored significantly higher than younger students on the composite section of the test. Therefore, hypothesis 1C was retained.

Hypothesis Two

Hypothesis two stated that second grade girls (older and younger) would score significantly higher on reading scores, total math scores, and composite scores of the Iowa Test of Basic Skills (Form 7, Level 7) than would second grade boys (older and younger). For purposes of analysis
TABLE I

SUM OF SQUARES, DEGREES OF FREEDOM, MEAN SQUARES, F, AND SIGNIFICANCE OF F FOR READING SCORES FOR OLDER AND YOUNGER STUDENTS ON THE IOWA TEST OF BASIC SKILLS

<table>
<thead>
<tr>
<th>Source</th>
<th>D.F.</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F</th>
<th>Significance Of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>1</td>
<td>3.3550</td>
<td>3.3550</td>
<td>4.2035</td>
<td>.0407</td>
</tr>
<tr>
<td>Within Groups</td>
<td>801</td>
<td>639.3149</td>
<td>.7981</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>802</td>
<td>642.6699</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source</td>
<td>D.F.</td>
<td>Sum of Squares</td>
<td>Mean Squares</td>
<td>F</td>
<td>Significance of F</td>
</tr>
<tr>
<td>------------------</td>
<td>------</td>
<td>----------------</td>
<td>--------------</td>
<td>------</td>
<td>------------------</td>
</tr>
<tr>
<td>Between Groups</td>
<td>1</td>
<td>10.3446</td>
<td>10.3446</td>
<td>28.6432</td>
<td>.0001</td>
</tr>
<tr>
<td>Within Groups</td>
<td>801</td>
<td>289.2845</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>802</td>
<td>299.6291</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE III

SUM OF SQUARES, DEGREES OF FREEDOM, MEAN SQUARES, F, AND SIGNIFICANCE OF F
FOR THE COMPOSITE SCORES FOR OLDER AND YOUNGER STUDENTS ON
THE IOWA TEST OF BASIC SKILLS

<table>
<thead>
<tr>
<th>Source</th>
<th>D.F.</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F</th>
<th>Significance of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>1</td>
<td>3.6128</td>
<td>3.6128</td>
<td>6.5289</td>
<td>.0108</td>
</tr>
<tr>
<td>Within Groups</td>
<td>801</td>
<td>443.2297</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>802</td>
<td>446.8424</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
and reporting, this hypothesis was divided into three sub-
hypotheses. Hypothesis 2A related sex to reading scores. 
As shown in Table IV, the one way analysis of variance 
indicated an F of 10.50, which was significant beyond the 
.001 level. Girls scored significantly higher than boys in 
reading. Therefore, hypothesis 2A was retained.

Hypothesis 2B related sex to total math scores. As 
shown in Table V, the one way analysis of variance indicated 
an F of 3.13, which was not significant. Girls did not 
score significantly higher than boys in math. Therefore, 
hypothesis 2B was rejected.

Hypothesis 2C related sex to composite scores. As 
shown in Table VI, the one way analysis of variance 
indicated an F of 3.31, which was not significant. Girls 
did not score significantly higher than boys on the 
composite section of the test. Therefore, hypothesis 2C 
was rejected.

Hypothesis Three

Hypothesis three stated that second grade girls whose 
birth dates were between September and November in 1975 
(older girls) would score significantly higher on reading 
scores, total math scores, and composite scores of the Iowa 
Test of Basic Skills (Form 7, Level 7) than second grade 
boys who were the same age. For purposes of analysis and 
reporting, the hypothesis was divided into three
### TABLE IV

SUM OF SQUARES, DEGREES OF FREEDOM, MEAN SQUARES, F, AND SIGNIFICANCE OF F FOR READING SCORES FOR BOYS AND GIRLS ON THE IOWA TEST OF BASIC SKILLS

<table>
<thead>
<tr>
<th>Source</th>
<th>D.F.</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F</th>
<th>Significance of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>1</td>
<td>8.3125</td>
<td>8.3125</td>
<td>10.4962</td>
<td>.0012</td>
</tr>
<tr>
<td>Within Groups</td>
<td>801</td>
<td>634.3574</td>
<td>.7920</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>802</td>
<td>642.6699</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**TABLE V**

**SUM OF SQUARES, DEGREES OF FREEDOM, MEAN SQUARES, F, AND SIGNIFICANCE OF F FOR TOTAL MATH SCORES FOR BOYS AND GIRLS ON THE IOWA TEST OF BASIC SKILLS**

<table>
<thead>
<tr>
<th>Source</th>
<th>D.F.</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F</th>
<th>Significance of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>1</td>
<td>1.1646</td>
<td>1.1646</td>
<td>3.1254</td>
<td>.0775</td>
</tr>
<tr>
<td>Within Groups</td>
<td>801</td>
<td>298.4645</td>
<td>.3726</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>802</td>
<td>299.6291</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


TABLE VI

SUM OF SQUARES, DEGREES OF FREEDOM, MEAN SQUARES, F, AND SIGNIFICANCE OF F
FOR COMPOSITE SCORES FOR BOYS AND GIRLS ON THE
IOWA TEST OF BASIC SKILLS

<table>
<thead>
<tr>
<th>Source</th>
<th>D.F.</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F</th>
<th>Significance of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>1</td>
<td>1.8371</td>
<td>1.8371</td>
<td>3.3068</td>
<td>.0694</td>
</tr>
<tr>
<td>Within Groups</td>
<td>801</td>
<td>445.0053</td>
<td>.5556</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>802</td>
<td>446.8424</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

sub-hypotheses. Hypothesis 3A related age and sex to reading scores. As presented in Table VII, one way analysis of variance indicated an F of 5.13 which was significant beyond the .01 level. The Scheffe multiple comparison method was used for each pair of group means to determine where the significance occurred. The group means for the ITBS reading scores are presented in Table VIII. Older girls scored significantly higher than older boys in reading. Hypothesis 3A, therefore, was retained.

Hypothesis 3B related sex and age to total math scores. One way analysis of variance indicated an F of 11.05 which was significant beyond the .01 level as presented in Table IX. The group means for the ITBS total math scores are presented in Table X. Using the Scheffe method, it was found that older girls did not score significantly higher in math than older boys. Therefore, hypothesis 3B was rejected.

Hypothesis 3C related age and sex to the composite scores. As presented in Table XI, one way analysis of variance indicated an F of 3.29 which was significant beyond the .05 level. The Scheffe multiple comparison method was used to determine where the significance occurred. Table XII indicates that older girls did not score significantly higher than boys on the composite score. Therefore, hypothesis 3C was rejected.
TABLE VII

SUM OF SQUARES, DEGREES OF FREEDOM, MEAN SQUARES, F, AND SIGNIFICANCE OF F FOR READING SUBTEST FOR BOYS AND GIRLS ON THE IOWA TEST OF BASIC SKILLS

<table>
<thead>
<tr>
<th>Source</th>
<th>D.F.</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F</th>
<th>Significance of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>3</td>
<td>12.1367</td>
<td>4.0456</td>
<td>5.1264</td>
<td>.0016</td>
</tr>
<tr>
<td>Within Groups</td>
<td>799</td>
<td>630.5332</td>
<td>.7892</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>802</td>
<td>642.6699</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE VIII
MEANS FOR YOUNGER BOYS AND GIRLS AND OLDER BOYS AND GIRLS ON THE
READING SUBTEST OF THE IOWA TEST OF BASIC SKILLS

<table>
<thead>
<tr>
<th>Group</th>
<th>Grade Equivalent Means</th>
<th>Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Younger Boys</td>
</tr>
<tr>
<td>Younger Boys</td>
<td>2.5591</td>
<td></td>
</tr>
<tr>
<td>Younger Girls</td>
<td>2.7139</td>
<td></td>
</tr>
<tr>
<td>Older Boys</td>
<td>2.6392</td>
<td>*</td>
</tr>
<tr>
<td>Older Girls</td>
<td>2.8920</td>
<td>*</td>
</tr>
</tbody>
</table>

*Indicates significant difference at or beyond the .05 level.
<table>
<thead>
<tr>
<th>Source</th>
<th>D.F.</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F</th>
<th>Significance of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>3</td>
<td>11.9383</td>
<td>3.9794</td>
<td>11.0520</td>
<td>.0001</td>
</tr>
<tr>
<td>Within Groups</td>
<td>799</td>
<td>287.6909</td>
<td>.3601</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>802</td>
<td>299.6291</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE X
MEANS FOR YOUNGER BOYS AND GIRLS AND OLDER BOYS AND GIRLS ON THE
TOTAL MATH SCORES OF THE IOWA TEST OF BASIC SKILLS

<table>
<thead>
<tr>
<th>Group</th>
<th>Grade Equivalent Means</th>
<th>Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Younger Boys</td>
</tr>
<tr>
<td>Younger Boys</td>
<td>2.1227</td>
<td></td>
</tr>
<tr>
<td>Younger Girls</td>
<td>2.0916</td>
<td>*</td>
</tr>
<tr>
<td>Older Boys</td>
<td>2.3955</td>
<td>*</td>
</tr>
<tr>
<td>Older Girls</td>
<td>2.2729</td>
<td>*</td>
</tr>
</tbody>
</table>

*Indicates significant difference at or beyond the .05 level.
TABLE XI

SUM OF SQUARES, DEGREES OF FREEDOM, MEAN SQUARES, F, AND SIGNIFICANCE OF F FOR THE COMPOSITE SCORES FOR BOYS AND GIRLS ON THE IOWA TEST OF BASIC SKILLS

<table>
<thead>
<tr>
<th>Source</th>
<th>D.F.</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F</th>
<th>Significance of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>3</td>
<td>5.4519</td>
<td>1.8173</td>
<td>3.2897</td>
<td>.0202</td>
</tr>
<tr>
<td>Within Groups</td>
<td>799</td>
<td>441.3905</td>
<td>.5524</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>802</td>
<td>446.8424</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE XII

MEANS FOR YOUNGER BOYS AND GIRLS AND OLDER BOYS AND GIRLS ON THE COMPOSITE SCORES OF THE IOWA TEST OF BASIC SKILLS

<table>
<thead>
<tr>
<th>Mean</th>
<th>Grade Equivalent Mean</th>
<th>Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Younger Boys</td>
</tr>
<tr>
<td>Younger Boys</td>
<td>2.3975</td>
<td></td>
</tr>
<tr>
<td>Younger Girls</td>
<td>2.4866</td>
<td></td>
</tr>
<tr>
<td>Older Boys</td>
<td>2.5251</td>
<td></td>
</tr>
<tr>
<td>Older Girls</td>
<td>2.6271</td>
<td></td>
</tr>
</tbody>
</table>

*Indicates significant difference at or beyond the .05 level.
Hypothesis Four

Hypothesis four stated that younger girls in second grade would score significantly higher on reading scores, total math scores, and composite scores on the Iowa Test of Basic Skills (Form 7, Level 7) than younger boys in second grade. For purposes of analysis and reporting, this hypothesis was divided into three sub-hypotheses. Hypothesis 4A related sex and age to reading scores. As presented in Table VII, one way analysis of variance indicated a significant F beyond the .01 level. The Scheffe multiple comparison method was used for each pair of group means. The group means for the ITBS reading scores are presented in Table VIII. This indicates that younger girls did not score significantly higher in reading than younger boys. Therefore, hypothesis 4A was rejected.

Hypothesis 4B related sex and age to total math scores. One way analysis of variance indicated a significant F beyond the .01 level as presented in Table IX. The group means for total math scores are presented in Table X. The Scheffe multiple comparison indicated that younger girls did not score significantly higher in math than younger boys. Therefore, hypothesis 4B was rejected.

Hypothesis 4C related sex and age to composite scores. As presented in Table XI, the one way analysis of variance indicated an F significant beyond the .05 level. Group
Hypothesis Five

Hypothesis five stated that older girls in second grade would score significantly higher on reading scores, total math scores, and composite scores of the Iowa Test of Basic Skills (Form 7, Level 7) than younger girls in second grade. For purposes of analysis and reporting, the hypothesis was divided into three sub-hypotheses. Hypothesis 5A related sex and age to reading scores. As presented in Table VII, one way analysis of variance indicated an F significant beyond the .01 level. The Scheffe multiple comparison method was used for each pair of group means. The group means for the ITBS reading scores are presented in Table VIII. Older girls did not score significantly higher in reading than younger girls. Hypothesis 5A, therefore, was rejected.

Hypothesis 5B related sex and age to total math scores. One way analysis of variance indicated an F significant beyond the .01 level as presented in Table IX. The group means for the ITBS math scores are presented in Table X. Using the Scheffe multiple comparison method it was found
that older girls did score significantly higher in math than younger girls. Therefore, hypothesis 5B was retained.

Hypothesis 5C related sex and age to the composite scores. As presented in Table XI, one way analysis of variance indicated an F significant beyond the .05 level. The group means for the ITBS composite scores are presented in Table XII. The Scheffe multiple comparison method indicated that the older girls did not score significantly higher on the composite score than younger girls. Hypothesis 5C was rejected.

Hypothesis Six

Hypothesis six stated that older boys in second grade would score significantly higher on reading scores, total math scores, and composite scores of the Iowa Test of Basic Skills (Form 7, Level 7) than younger boys in second grade. For purposes of analysis and reporting, the hypothesis was divided into three sub-hypotheses. Hypothesis 6A related sex and age to reading. As presented in Table VII, one way analysis of variance indicated an F significant beyond the .01 level. The Scheffe multiple comparison method was used for each pair of group means. The group means for the ITBS reading scores are presented in Table VIII. The older boys did not score significantly higher in reading than the younger boys. Hypothesis 6A, therefore, was rejected.
Hypothesis 6B related sex and age to total math scores. One way analysis of variance indicated an F significant beyond the .01 level as presented in Table IX. The group means for the total math scores are presented in Table X. The Scheffe multiple comparison method indicated that the older boys did score significantly higher in math than the younger boys. Therefore, hypothesis 6B was retained.

Hypothesis 6C related sex and age to composite scores. As presented in Table XI, one way analysis of variance indicated an F significant beyond the .05 level. Group means for the composite scores are presented in Table XII. The Scheffe multiple comparison indicated that older boys did not score significantly higher on the composite score than younger boys. Hypothesis 6C was, therefore, rejected.

Hypothesis Seven

Hypothesis seven stated that older girls in second grade would score significantly higher on reading scores, total math scores, and composite scores of the Iowa Test of Basic Skills (Form 7, Level 7) than younger boys in second grade. For purposes of analysis and reporting, the hypothesis was divided into three sub-hypotheses. Hypothesis 7A related sex and age to reading scores. As presented in Table VII, one way analysis of variance indicated an F significant beyond the .01 level. The Scheffe multiple comparison method was used for each pair of group means.
The group means for the ITBS reading scores are presented in Table VIII. This indicated that older girls scored significantly higher in reading than younger boys. Therefore, hypothesis 7A was retained.

Hypothesis 7B related sex and age to total math scores. One way analysis of variance indicated an F significant beyond the .01 level as presented in Table IX. The group means for total math are presented in Table X. The Scheffe multiple comparison method indicated that older girls scored significantly higher in math than younger boys. Therefore, hypothesis 7B was retained.

Hypothesis 7C related sex and age to composite scores. As presented in Table XI, one way analysis of variance indicated an F significant beyond the .05 level. Group means for composite scores are presented in Table XII. The Scheffe multiple comparison indicated that older girls scored significantly higher on the composite section of the ITBS than younger boys. Hypothesis 7C was retained.

Hypothesis Eight

Hypothesis eight stated that older boys in second grade would not score significantly higher on reading scores, total math scores, and composite scores of the Iowa Test of Basic Skills (Form 7, Level 7) than would younger girls in second grade. For purposes of analysis and reporting, the hypothesis was divided into three
sub-hypotheses. Hypothesis 8A related sex and age to reading scores. As presented in Table VII, one way analysis of variance indicated an $F$ significant beyond the .01 level. The Scheffe multiple comparison method was used for each pair of group means. The group means for the ITBS reading scores are presented in Table VIII. Older boys did not score significantly higher in reading than younger girls. Hypothesis 8A, therefore, was retained.

Hypothesis 8B related sex and age to total math scores. One way analysis of variance indicated an $F$ significant beyond the .01 level as presented in Table IX. The group means for the ITBS total math scores are presented in Table X. The Scheffe multiple comparison method indicated that older boys scored significantly higher than younger girls. Therefore, hypothesis 8B was rejected.

Hypothesis 8C related sex and age to the composite scores. As presented in Table XI, one way analysis of variance indicated an $F$ significant beyond the .05 level. The group means for the ITBS composite scores are presented in Table XII. The Scheffe multiple comparison method indicated that older boys did not score higher than younger girls on the composite section of the test. Therefore, hypothesis 8C was retained.
Hypothesis Nine

Hypothesis nine stated that of those children retained in first grade, there would be more younger boys than any other group.

A chi square test of independence was the statistical procedure used to test the null hypothesis that there was no relationship between sex and age in the number of children retained in first grade. Table XIII presents the numbers retained in first grade by sex and age.

**TABLE XIII**

**NUMBER OF STUDENTS RETAINED IN FIRST GRADE**

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Young</th>
<th>Old</th>
<th>Total Sex Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boys</td>
<td>94</td>
<td>21</td>
<td>115</td>
</tr>
<tr>
<td>Girls</td>
<td>48</td>
<td>12</td>
<td>60</td>
</tr>
<tr>
<td>Total Age Group</td>
<td>142</td>
<td>33</td>
<td>175</td>
</tr>
</tbody>
</table>

The critical region for the .05 level of significance starts at 3.84. The chi square obtained from the numbers in Table XIII was calculated as .0779. Therefore, the null hypothesis was retained. There was no relationship shown between sex and age for those children retained in first grade.
Teacher Grades as Related to Sex and Age of Students

It was difficult to make any generalizations from teacher grades because of the grading system itself and the subjectivity involved. When grades are given that do not have a numerical base, such as E, S, and U, a great deal of subjectivity is involved. Each teacher is an individual with different philosophies, backgrounds, and expectations. All of these variables enter into the teacher's grade-making decisions.

Within the school district used in this study there were thirty-one elementary schools. Among the thirty-one elementary schools there were several different approaches to organization. There were those that made use of ability grouping in reading and math for the entire grade level. Some grouped the entire grade level, but only for reading. Other schools employed no ability grouping across the entire grade level. Each class was entirely self-contained and any ability grouping was left to the discretion of each individual teacher. The continuity and consistency of ability grouping was probably better within each individual elementary school. However, differences did exist even within schools.

The diversity of the neighborhood schools also had an impact on the grades and ability levels. A child in a below grade level group in one school could possibly have been in
an average or above average group at another school in a different section of town.

All of these influences, including different methods of ability grouping, different philosophies, different standards, different backgrounds, completely subjective grades without a numerical base, and diverse neighborhoods, had an impact on grades given. Therefore, grades from only one elementary school in a middle to upper middle class neighborhood were used in this study.

In order to examine teacher grades as related to sex and age of students, yearly averages were obtained from one elementary school for those students who fit into one of the following four categories:

1. Older boys (September to November birth dates),
2. Older girls (September to November birth dates),
3. Younger boys (June to August birth dates), and
4. Younger girls (June to August birth dates).

This group consisted of thirty-three children. There were six younger boys, seven younger girls, six older boys, and fourteen older girls. A reading grade, math grade, and ability group for each child are presented in Table XIV. The grades were the yearly average for first grade.

Grades of E, S+, S, S-, N, and U (excellent, high average, average, low average, needs improvement, and unsatisfactory) were given over an ability level. The ability levels were indicated by numbers. The number one
<table>
<thead>
<tr>
<th>Students</th>
<th>Grades</th>
<th></th>
<th>Ability Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reading</td>
<td>Math</td>
<td></td>
</tr>
<tr>
<td><strong>Young Boys:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adam</td>
<td>E/1</td>
<td>E/2</td>
<td>HA</td>
</tr>
<tr>
<td>Jason</td>
<td>S/2</td>
<td>S/2</td>
<td>LA</td>
</tr>
<tr>
<td>B. J.</td>
<td>a</td>
<td>a</td>
<td>LA</td>
</tr>
<tr>
<td>Tom</td>
<td>a</td>
<td>a</td>
<td>L</td>
</tr>
<tr>
<td>Jeremy</td>
<td>S+/2</td>
<td>S+/2</td>
<td>L</td>
</tr>
<tr>
<td>Eric</td>
<td>S/2</td>
<td>S/2</td>
<td>LA</td>
</tr>
<tr>
<td><strong>Young Girls:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kelly</td>
<td>S/2</td>
<td>S+/2</td>
<td>LA</td>
</tr>
<tr>
<td>Jennifer</td>
<td>S+/1</td>
<td>S+/2</td>
<td>H</td>
</tr>
<tr>
<td>Ginny</td>
<td>S/2</td>
<td>S/2</td>
<td>LA</td>
</tr>
<tr>
<td>Amanda</td>
<td>S-/2</td>
<td>S-/2</td>
<td>L</td>
</tr>
<tr>
<td>Gretchen</td>
<td>S+/2</td>
<td>S-/2</td>
<td>LA</td>
</tr>
<tr>
<td>Stephanie</td>
<td>S-/2</td>
<td>S/3</td>
<td>L</td>
</tr>
<tr>
<td>Nicole</td>
<td>S/3</td>
<td>S/3</td>
<td>L</td>
</tr>
<tr>
<td><strong>Old Boys:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thomas</td>
<td>S/3</td>
<td>S/3</td>
<td>L</td>
</tr>
<tr>
<td>Jeremy</td>
<td>S/2</td>
<td>S+/2</td>
<td>LA</td>
</tr>
<tr>
<td>Mark</td>
<td>S-/2</td>
<td>S/2</td>
<td>LA</td>
</tr>
<tr>
<td>Lawrence</td>
<td>S+/1</td>
<td>S+/2</td>
<td>H</td>
</tr>
</tbody>
</table>
TABLE XIV--Continued

<table>
<thead>
<tr>
<th>Students</th>
<th>Grades</th>
<th>Ability Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reading</td>
<td>Math</td>
</tr>
<tr>
<td>Old Boys:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steven</td>
<td>S+/1</td>
<td>S+/2</td>
</tr>
<tr>
<td>James</td>
<td>E/1</td>
<td>S+/2</td>
</tr>
<tr>
<td>Old Girls:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shannon</td>
<td>E/1</td>
<td>S+/2</td>
</tr>
<tr>
<td>Jenna</td>
<td>S+/2</td>
<td>S/2</td>
</tr>
<tr>
<td>Alyson</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>Kristine</td>
<td>S+/1</td>
<td>S+/2</td>
</tr>
<tr>
<td>Kim</td>
<td>S+/1</td>
<td>E/2</td>
</tr>
<tr>
<td>Michelle</td>
<td>S+/1</td>
<td>S+/2</td>
</tr>
<tr>
<td>Tricia</td>
<td>S/2</td>
<td>S/2</td>
</tr>
<tr>
<td>Cathy</td>
<td>S/2</td>
<td>S-/2</td>
</tr>
<tr>
<td>Jennifer</td>
<td>E/1</td>
<td>S+/2</td>
</tr>
<tr>
<td>Laura</td>
<td>S+/1</td>
<td>S+/2</td>
</tr>
<tr>
<td>Nancy</td>
<td>S+/1</td>
<td>S+/2</td>
</tr>
<tr>
<td>Sherry</td>
<td>E/1</td>
<td>E/2</td>
</tr>
<tr>
<td>Barbie</td>
<td>S/3</td>
<td>S/3</td>
</tr>
<tr>
<td>Nicole</td>
<td>S/2</td>
<td>S/2</td>
</tr>
</tbody>
</table>

Note. H = high; HA = high average; A = average; LA = low average; L = low; E = excellent; S, S+, S- = degrees of satisfactory; N = needs improvement; U = unsatisfactory; 1 = above grade level; 2 = on grade level; 3 = below grade level.

aGrades were not available for these students.
indicated above grade level, two was on grade level, and three was below grade level. For example, a student with a grade of S/3 was making satisfactory progress on below grade level material.

The information in Table XIV shows that only two of thirteen younger children (boys and girls) were placed in an average or higher ability group. From the twenty older children (boys and girls), only five were placed in a below average ability group, the majority of these being boys. These facts seem to correlate with the previous findings in this study, such as: older girls generally do better overall, especially in reading; older boys do better in math (based on ITBS scores); and younger boys have more problems overall than any other group.

The above grade level and the below grade level groups were the easiest to identify. The on grade level group usually included students in low average, average, and high average ability groups. Therefore, in order to interpret teacher grades, it was necessary to know how each child was ability grouped in relation to the other children in the class.

Of the six younger boys listed, only one was placed in a high average group. Two of these boys were retained in second grade. It had been requested that one of the two be retained in first grade, but the parents refused at that time.
The younger girls and the older boys were the most similar of the groups. The children from these groups appeared in every ability group, with the majority being on grade level.

Only two of the fourteen older girls were placed in low or low average ability groups. This correlates with the analysis of variance results which indicated a significant difference between age on all tests and in sex on the reading test.

Summary of Findings

The results of this study provide evidence that age at entrance to school does have an impact on achievement. In the area of reading, sex also proved to be a significant indicator of success. The widest range of difference existed between older girls and younger boys, favoring the girls.

In each section of the ITBS that was analyzed, older students scored higher. Older girls scored higher in reading and the composite sections. Older boys scored higher in math. Younger boys made the lowest scores in reading and composite sections. Younger girls scored lowest in math.

In order to include those children retained in first grade who were the same ages as the second graders in the study, a separate analysis using chi square was performed.
The results of these tests indicated that the largest group of children retained in first grade was young boys.

There did seem to be a correlation between first grade year end teacher grades and the analysis of variance results which indicated a significant difference between age on all tests and in sex on the reading test. However, many factors influence teacher grades, such as philosophies, expectations, grouping methods, and diversity of neighborhoods. Therefore, comparison with another elementary school in the same district could yield different results.
The present research study was an investigation to compare achievement of boys and girls in second grade who were seven years old in June, July, and August of 1983 to the boys and girls in second grade who were eight years old in September, October, and November of 1983. The students were tested using the Iowa Test of Basic Skills (Form 7, Level 7) in September of 1983 using the following areas: reading, total math, and composite scores. The study also looked at the correlation of sex and age of students who had been retained in first grade. A comparison of teacher grades to standardized test scores and ability grouping was also presented.

The final list of students in the study was taken from the entire first and second grade population of elementary schools in Garland, Texas. A computer list that contained sex and birth dates of all first and second grade students was generated. Second grade students chosen for the study had birth dates in June, July, and August in 1976 and September, October, and November in 1975. These students were then divided into four sections by sex and age. First grade students in the study were those who by chronological
age should have been in second grade. This was determined by birth dates. There was a total of 803 second grade students and 175 first grade students in the study.

The Iowa Test of Basic Skills (Form 7, Level 7) was the instrument employed in the study. This test was administered by the second grade classroom teachers in the district. The test was computer scored.

The statistical procedure applied to the test results was one way analysis of variance. Multiple comparisons using the Scheffe method, where appropriate, were used. Chi square was the statistical procedure used in the analysis of the students who were retained in first grade. A chi square test of independence was conducted to determine if interaction was indicated.

Findings

The major findings resulting from analysis of statistical data presented in this study were as follows:

1. Older students scored significantly higher than younger students on all three tests.

2. Second grade girls scored significantly higher than second grade boys in reading, but not in total math or on the composite score.

3. Older girls scored significantly higher than all boys in reading, but not in total math or on the composite score.
4. Younger girls did not score significantly higher than younger boys in reading, total math, or on the composite score.

5. Older girls scored significantly higher than younger girls in total math, but not in reading or on the composite score.

6. Older boys scored significantly higher than younger boys in total math, but not in reading or on the composite score.

7. Older girls scored significantly higher than younger boys in reading and on the composite score.

8. Older boys scored significantly higher than younger girls in total math, but not in reading or on the composite score.

9. There were significantly more younger boys retained in first grade than any other group.

10. Reading ability groups in the first grade tended to correlate with achievement test scores in second grade.

Discussion of Findings

Findings indicated significant difference between younger boys and older girls in reading and on the composite scores. Older girls scored significantly higher than younger boys in the sections of the test that required reading.
The results that showed girls scoring significantly higher in reading than boys in this study support previous research. Girls seem to have an inherited language advantage and as a group achieve better in reading and have fewer problems in school (Ames, 1972; Dechant, 1978).

On the math test there was a significant difference between (1) older boys and younger girls and boys and (2) older girls and the younger girls. The older children scored significantly higher in math. The review of literature indicated no research relating sex and age to achievement in math.

The composite scores showed significant difference only between older girls and younger boys. Since the composite score included reading and math, the differences in abilities of either sex for one particular subject would be balanced. Hence, the only significant difference was between the two extremes (older girls and younger boys).

These findings correlate with research done by the Gesell Institute. The Gesell Developmental Test included sections that require language ability plus visual and spatial ability. Since girls tend to score higher in language ability and boys tend to score higher in visual and spatial ability, this test balanced the differences between the sexes as did the composite score on the ITBS.

The chi square test calculated on the numbers of children retained in first grade indicated no relationship
between sex and age. However, a higher number of younger boys were retained than any other group. This supports the finding of Hall (1963). In his study he found that 801 elementary pupils had been retained at one grade level. Almost three-fourths of these were boys.

In looking at first grade year end teacher grades and reading ability groups from one school, it was found that the ability groups correlated with the findings of the analysis of variance that showed older girls score higher in reading and younger boys score lower. Because of the subjectivity and inconsistency of the teacher grades, it was difficult to use them as indicators of achievement. There are schools that use ability grouping for the entire grade level in reading and math and those that use ability grouping only for reading. There are also schools that are self-contained, with ability grouping varying from class to class. The difference in philosophies of teachers and the diversity of the neighborhood schools would also have an impact on the grades.

This study was significant in that it was unlike any research done to the present time. It not only compared age (older children and younger children) and sex (boys and girls), but also compared every possible combination involving age and sex. It also involved more than one subject area, which brought out the significance of older boys scoring higher in math.
Conclusions

Older children scored higher in all three areas measured. However, girls scored higher in reading. This may seem contradictory. However, it is not. Age was significant beyond the .05 level in reading, while girls scored significantly higher beyond the .001 level. This difference is explained by the extremes in means for younger boys and older girls. Since first grade curriculum emphasizes reading, this gives girls a definite advantage over boys, who scored higher in math. This indicates a need for restructuring curriculum to meet the needs of both sexes.

Younger boys in second grade scored the lowest in all areas tested, except math. These scores would have been even more significant if the large number of younger boys who were the same age, but were retained in first grade, could have been included in the second grade testing. These findings indicate that total developmental age is the important factor when considering entrance for school.

Recommendations

Because of the significant results of this study, a closer look should be taken at the requirements for school entrance. Historically school admission has been based on chronological age. Evidence accumulated since that time has challenged the practice of a fixed age for school admission (Barrett, 1956; Durrell, 1958; Dykstra, 1966; Rowland & Nelson, 1959).
No child fits all the norms for his or her age. Each will have some behavior typical of younger children and some typical of older children (Levenson, 1977). Young children do not learn in the same way as older children and adults (Elkind, 1986). Because of this difference, chronological age is not a meaningful predictor of school success. Children exactly six years old will vary as much as 2.3 years in mental age (Hedges, 1977).

There is no single criterion for the determination of school entrance that adequately predicts school success. The factors that must be considered include chronological age, mental age, physical maturity, emotional and social maturity, behavior age, and sex (Braga, 1971). These factors were shown to be important in research conducted in 1972 by Ames showing that five times as many boys as girls were taken by their parents for clinic help. Research also found that young children are more likely to be referred for academic problems than older children (DiPasquale, 1980).

The public at this time is demanding more from education and accountability from educators. This is causing unrest in the education field and a variety of proposed reforms. Now would be an advantageous time to propose changes in school entrance procedures to a public that seems ready for some changes.

One major problem that will have to be confronted and will be the major obstacle to any proposal of change is
tradition. Children have begun first grade at six years of age since the beginning of this century. Adult relatives and friends of children associate grade level and age. If the two do not correlate by traditional standards, it is usually assumed there is some kind of problem with the child. It will be hard to convince the adult community that age and grade in school are not necessarily the same for all children.

Another problem confronting any proposed change or innovative alternatives in public education will be money. Most changes will require either a shift in budget appropriations or additional tax revenues in order to support the change. Again, the public must be convinced the money is for a worthwhile cause and is in the best interest of the entire community.

No matter how many educators agree on the course of action that should be taken, nothing can be effectively accomplished until the general public also agrees and supports the efforts of educators. Therefore, before a major change can successfully take place, the public needs to be informed as to the reasons for the change, the estimated costs involved, and the expected results.

A good way to educate the general public would be to start with parents who are directly involved. This is the foundation upon which change will be built. The movement will profit from enthusiastic parents who are convinced that change is best for the children. Outside speakers, films,
and pamphlets will help educate the public. Parent education should be repeated every year for several years.

By keeping the possible problems and solutions to these problems in mind, some of the following alternatives to current practices could possibly be implemented.

Research shows that there are different rates of maturation for boys and girls, with girls generally maturing at a faster pace, except in visual and spatial skills (Ames, 1972). An alternative to the current method of school admission that would help alleviate this difference and add no extra cost to the budget would be to raise the entrance age for boys by six months. However, this alternative considers only two broad groups, boys and girls. There would still be girls pushed on who were not ready, especially in math, and boys held back unnecessarily. The popularity of this alternative would be the fact that an effort was being made to address the problem and that no increase in funds would be needed.

Another alternative would be to continue the current practice of school admission, but provide an extended kindergarten or a transitional grade between kindergarten and first grade for those children who are not developmentally ready for the rigors of a more structured learning situation. Children should be able to sustain emotionally, socially, physically, and intellectually in the school setting. The extra cost involved would be the possible addition of
teachers for this level and the fact that these children would be in school longer (beyond eighteen).

In order to identify these children, the kindergarten teachers could recommend developmental testing for those children they thought would have difficulty in a traditional first grade situation. An advantage to this type of program would hopefully result in fewer retentions. The children would be perceived as progressing to another level instead of having the label of a failure. It would also be an effort on the part of the schools to recognize the difference in the development of children and to design a program to meet these needs.

An alternative that would require more restructuring within the school would be the nongraded school. In this situation each child would be permitted to progress at his own rate. This would mean that teachers as well as students would be crossing traditional grade level lines. There would have to be a great deal of cooperation among teachers in this situation in order to be sure that each child progressed when necessary. There should be no pressure put on children to complete certain curriculum at a predetermined time. This alternative would probably be the most difficult to implement and maintain.

Another alternative would be to establish a screening process using a developmental test, such as the one used at the Gesell Institute, as one of the screening devices.
Every child entering school should be placed through this screening process, preferably in the spring or early summer before entering school in the fall. One problem that would have to be faced would be the training of personnel to administer the developmental tests. This would probably involve a considerable amount of money in the initial stages. However, it would not be necessary to hire additional staff. Kindergarten and first grade teachers who were interested could be trained and employed for eleven months instead of ten.

Another problem to consider would be what to do with those children who indicate through the screening process that they are not developmentally ready for kindergarten. It would have to be decided if a public pre-kindergarten class would be provided for those children. Again, budget problems would influence the decision. Another year at home or private pre-kindergarten would also be alternatives.

There would also be the question of students entering at different levels of the elementary school from other districts that might not make use of developmental screening. A decision would need to be made as to whether or not these students would be tested. Another dilemma would be, if the child were tested and found developmentally inferior to the level in which he was placed, what would be the proper procedure for placement?

Instead of screening all children as they enter kindergarten, they could all be tested at the end of kindergarten
This differs from screening only those recommended by the kindergarten teachers as mentioned earlier. This would again require the addition of extended kindergarten classes or pre-first grade classes.

Another alternative would be to restructure first grade curriculum. The curriculum should be introduced according to the developmental level of the students. Possibly, first grade curriculum should be changed so that some students have a heavy math curriculum at the beginning of the year (older boys in particular), and some students have a heavy reading curriculum at the beginning of the year. The emphasis could be changed as the needs of the students change.

In each alternative mentioned there are foreseeable problems, the most significant being tradition and money. Some are more practical as a starting point for change than others and would be more easily implemented.

The significant results of this study emphasize the need for providing an alternative to using chronological age as the only criterion for admission to public school. Because of the numbers of children involved, any procedure that involves considering the individual characteristics of each child will require more time, effort, and money. Educators, the public, and the governing establishment will have to decide if the benefits warrant all the extras that would be demanded.
A longitudinal study of the students in this study would be justified to determine if the differences would continue to be significant as the students get older. Also, replication of this study in other areas of the country would be justified to determine if the results are applicable in more than one school district.
CHAPTER BIBLIOGRAPHY


February 20, 1984

Ms. Sharon Good
1805 Caldwell
Garland, Texas 75041

Dear Ms. Good:

The research council has concluded its evaluation of your application to conduct a research study in the Garland Independent School District. It is with pleasure that I inform you that the council approved your study, "The Relationship of Sex and Age at Entrance to School on Second Grade Achievement."

You may begin data collection activities in our district beginning this spring, 1984. Notify my office in writing of the date that you will begin your data collection activity and the date on which it will be concluded. Dr. Barbara Mason will be your contact person for the district and will assist you by coordinating data collection. You should inform the subjects that your research project has been approved by the Planning, Research, and Evaluation Department.

Three copies of your application to conduct a research study may be picked up from the PRE office. If we do not hear from you within two weeks regarding this, these applications will be discarded. The remaining materials will be maintained in our files.

Speaking for the research council, I wish you the best of success in your research efforts and look forward to receiving a copy of your report. Upon completion, your report should be filed with the Planning, Research, and Evaluation Department. If I may assist in any way, please contact me.

Sincerely yours,

Michael W. Strozeski, Ph.D.
MWS/fs
BIBLIOGRAPHY

Books


**Articles**


Unpublished Materials


