

1961

A study of stanines as a methodology in the identification of the potentially superior high-school student

<https://hdl.handle.net/2144/27201>

Downloaded from DSpace Repository, DSpace Institution's institutional repository

Thesis
Fee, K.E.
1961

BOSTON UNIVERSITY
SCHOOL OF EDUCATION

Thesis

A STUDY OF
STANINES AS A METHODOLOGY IN THE IDENTIFICATION OF
THE POTENTIALLY SUPERIOR HIGH-SCHOOL STUDENT

Submitted by

Kathryn E. Fee
(Ed. B., Rhode Island College, 1937)

In Partial Fulfillment of Requirements for
the Degree of Master of Education

1961

First Reader: William C. Kvaraceus
Professor of Education

Second Reader: Lawrence Litwack
Assistant Professor of Education

Third Reader: June E. Holmes
Assistant Professor of Education

TABLE OF CONTENTS

CHAPTER	Page
I. THE PROBLEM.....	1
Statement of the Problem.....	1
Purpose of the study.....	2
Justification of the study.....	3
Scope of the study.....	4
Definition of the Terms Used.....	7
Summary.....	9
II. REVIEW OF RELATED LITERATURE.....	10
Ways of Expressing Test Scores.....	12
Grade equivalents.....	13
Age equivalents.....	14
Modal-age grade norms.....	14
Percentiles.....	15
Standard scores.....	16
Stanine scores.....	17
Stanine studies.....	19
Reports on Applications of Stanines.....	22
III. PROCEDURES.....	25
Student Population.....	25
Methodology.....	27
Source of the Data.....	27
Intelligence measures.....	28
Achievement measures.....	28
Description and Evaluation of the Measuring Instruments Used.....	30
California Short-Form Test of Mental Maturity.....	30
Otis Quick-Scoring Mental Ability Test.....	32

CHAPTER		Page
III.	Metropolitan Achievement Tests.....	32
	Stanford Achievement Tests.....	34
	Cooperative Achievement Tests.....	35
	Iowa Tests of Educational Development.....	36
	Summary of the Measuring Instruments.....	38
	Organization and Treatment of the Data.....	38
	Recording of the data.....	38
	Plotting of stanine charts.....	39
	Ability versus achievement.....	41
	Achievement versus achievement.....	43
	Ability versus ability.....	46
	Summary of Procedures.....	47
IV.	SUMMARY OF FINDINGS AND TENTATIVE CONCLUSIONS...	50
	Determination of Stanine Stability.....	50
	Stability Between Mental Ability Test Scores.	51
	Stability Between CTMM Scores and Achievement	52
	Stability at grade six.....	52
	Stability at grade seven.....	54
	Stability at grade nine.....	56
	Significant Deviations Between CTMM Scores and Achievement.....	56
	Deviations in reading.....	58
	Deviations in arithmetic reasoning.....	61
	Total deviations in reading and arithmetic reasoning.....	62
	Deviations in other achievement.....	63
	Total deviations in all subjects.....	63
	Summary of Stanine Stability Between CTMM Scores and Achievement.....	65
	Summary of the findings.....	65
	Implications of the findings.....	68
	Stability Between Otis Ability Scores and Achievement.....	69
	Stability at grade nine.....	69
	Stability at grades nine and ten.....	71

CHAPTER

Page

IV.	Stability at grade eleven.....	74
	Total stability at grades nine, ten, and eleven.....	76
	Different Students Involved in Significant Deviations Between Ability Scores and Achievement.....	77
	Deviations in reading.....	78
	Deviations in arithmetic reasoning.....	80
	Deviations in reading and arithmetic reasoning.....	81
	Deviations in English.....	81
	Deviations in other subjects.....	82
	Deviations in all subjects.....	82
	Summary of Stanine Stability Between Ability and Achievement.....	83
	Results of the findings.....	83
	Different students involved in significant deviations.....	86
	Implications of the findings.....	87
	Stability in Achievement Test Scores at Different Grade Levels.....	89
	Stability when grade-six achievement is used as a point of reference.....	89
	Stability when grade-seven achievement is used as a point of reference.....	93
	Stability when grade-nine achievement is used as a point of reference.....	101
	Stability in grades ten and eleven.....	105
	Significant Deviations in Achievement and Number of Different Students Involved At Different Grade Levels.....	107
	Deviations in reading.....	107
	Deviations in arithmetic reasoning.....	110
	Deviations in English.....	112
	Deviations in vocabulary.....	115
	Deviations in arithmetic computation and spelling.....	116
	Deviations in all subjects.....	117
	Total Stability in Achievement Test Scores At All Grade Levels.....	119
	Stability in reading.....	119

CHAPTER		Page
IV.	Stability in arithmetic reasoning.....	121
	Stability in English.....	122
	Stability in vocabulary.....	123
	Composite stability by subject.....	123
	Composite stability by achievement tests...	125
	Ability and Previous Achievement as Reference	
	Points in Stanine Stability.....	127
	Grade six to eleven.....	127
	Deviation patterns.....	127
	Total stability.....	130
	Sixth-Grade Ability and Achievement As	
	Criteria of High School Achievement.....	132
	Measures of high-school achievement.....	133
	Ability stanines as criteria.....	134
	Average achievement stanines as criteria...	137
	Predicting achievement.....	138
V.	CONCLUDING SUMMARY, IMPLICATIONS, LIMITATIONS, AND RECOMMENDATIONS.....	140
	Stability Between Mental Ability Test Scores.	140
	Stability Between Mental Ability and Achieve- ment Test Scores.....	141
	CTMM-achievement differentials.....	141
	Otis-achievement differentials.....	142
	Overachievement and underachievement.....	143
	Total stability.....	145
	Implications.....	145
	Stability Between Achievement Test Scores....	146
	Grade six to eleven.....	147
	Grade seven to eleven.....	147
	Grade nine to eleven.....	147
	Plus and minus achievement.....	148
	Total stability.....	150
	Implications.....	150
	Ability and Previous Achievement as Reference	
	Points.....	151
	Sixth-Grade Ability and Achievement as	
	Criteria of High-School Achievement.....	152
	Prediction of stability.....	153

CHAPTER

Page

V.	Prediction of superior achievement.....	153
	Evaluation of the Stanine Methodology.....	155
	Advantages.....	155
	Disadvantages.....	156
	Limitations of the Study.....	156
	Recommendations.....	156
	For the city of Providence.....	156
	For further study.....	157

APPENDIX	Page
SAMPLE STANINE TALLY CHART.....	160
SUMMARY TABLES OF STANINE VARIATIONS.....	161
Ability-Achievement.....	161
Achievement-Achievement.....	164
Composite.....	169
SUMMARY TABLE OF ACHIEVEMENT STANINES.....	170
SELECTED BIBLIOGRAPHY.....	172

LIST OF TABLES

Table	Page
1. Proportion of Cases, and National Modal Percentile Ranks Included in Each National Stanine Level.....	39
2. Tests Used in Computation of Ability-Achievement Stanine Differentials.....	42
3. Tests Used in the Computation of Achievement-Achievement Stanine Differentials.....	44
4. Stanine Variations in the Scores of 194 Students Between the California Short-Form Test of Mental Maturity-Grade Six, and the Otis Quick-Scoring Test of Mental Ability-Grade Nine.....	51
5. Stanine Stability in the Scores of 183 Students Between the California Short-Form Test of Mental Maturity and the Sub-tests of Reading and Arithmetic Problems of the Metropolitan Achievement Tests-Grade Six.....	53
6. Stanine Stability in the Scores of 194 Students Between the California Short-Form Test of Mental Maturity and each Sub-test of the Stanford Achievement Test in Grade Seven.....	55
7. Stanine Stability in the Scores of 194 Students Between the California Short-Form Test of Mental Maturity and Each Sub-test of the Stanford Achievement Test in Grade Nine.....	57
8. A Comparison of the Significant Deviations in the Achievement Scores of the Same Group of 194 Students As They Progressed from the Sixth Grade to the Ninth, When the CTMM Is Used as a Point of Reference.....	59
9. A Comparison of the Stability in the Scores of the Same 194 Students Between the California Short-Form Test of Mental Maturity and All Achievement Tests in Grades Six, Seven, and Nine	67

Table

Page

10.	Stanine Stability in the Scores of 200 Students Between the Otis Test of Mental Ability and Each Sub-test of the Stanford Achievement Test in Grade Nine.....	70
11.	Stanine Stability in the Scores of 200 Students Between the Otis Test of Mental Ability and the Cooperative Tests in Grades Nine and Ten...	72
12.	Stanine Variations in the Scores of 121 Selected Students Between the Otis Test of Mental Ability and the Iowa Tests of Educational Development in the Eleventh Grade.....	75
13.	A Comparison of the Stanine Stability in the Scores of the Same 200 Students Between the Otis Test of Mental Ability and All Achievement Tests in Grades Nine to Eleven.....	77
14.	Number of Different Students, in Group of 200 Students, Involved in Significant Deviations Between Ability and Achievement in Grades Six, Seven, Nine, and Ten.....	79
15.	A Comparison of the Total Stanine Variations in All Ability-Achievement Scores of the Same 200 Students As They Progressed from the Sixth Grade to the Eleventh.....	85
16.	A Comparison of the Stanine Stability in the Reading Performance of the Same Group of 184 Students in Grades Seven, Nine, Ten, and Eleven When Grade-Six Achievement Is Used as a Point of Reference.....	90
17.	A Comparison of the Stanine Stability of Test Scores in Arithmetic Reasoning of the Same Group of 182 Students When Grade-Six Achievement Is Used as a Point of Reference.....	91
18.	A Comparison of the Stanine Stability in the Composite Scores of Reading and Arithmetic Reasoning of the Same Group of 184 Students When Grade-Six Achievement Is Used as a Point of Reference.....	92

Table

Page

19.	A Comparison of the Stanine Stability in the Reading Performance of the Same Group of 196 Students in Grades Nine to Eleven When Seventh-Grade Achievement Is Used As a Point of Reference.....	94
20.	A Comparison of the Stanine Stability in the Test Scores in Arithmetic Reasoning of the Same Group of 193 Students When the Seventh-Grade Achievement Is Used As a Point of Reference....	95
21.	A Comparison of the English Performance of the Same Group of 183 Students in Grades Nine to Eleven When the Seventh-Grade Achievement Is Used As a Point of Reference.....	96
22.	A Comparison of the Stanine Stability in the Vocabulary Scores of the Same Group of 195 Students in Grades Nine to Eleven When Seventh-Grade Achievement Is Used As a Point of Reference.....	97
23.	A Comparison of the Performance of the Same Group of 195 Students in Arithmetic Computation and Spelling in the Ninth Grade When Seventh-Grade Achievement Is Used As a Point of Reference....	98
24.	A Comparison of the Stanine Stability in the Performance of the Same Group of 200 Students on All Subsequent Achievement Tests When Grade-Seven Achievement Is Used As a Point of Reference.....	100
25.	A Comparison of the Stanine Stability in the Reading Performance of the Same Group of 200 Students in Grades Ten and Eleven When Ninth-Grade Achievement Is Used As a Point of Reference.....	101
26.	A Comparison of the Stanine Stability in the Test Scores in Arithmetic Reasoning of the Same Group of 193 Students in Grades Nine A and Eleven When Nine B Achievement Is Used As a Point of Reference.....	102

Table	Page
27. A Comparison of the Stanine Stability in the Test Scores in English of the Same Group of 200 Students in Grades Nine to Eleven When Ninth-Grade Achievement Is Used As a Point of Reference.....	103
28. A Comparison of the Stanine Stability in the Performance of the Same Group of 200 Students on All Subsequent Achievement Tests When Grade-Nine Achievement Is Used As a Point of Reference.....	104
29. A Comparison of the Stanine Stability in the Achievement Test Scores of the Same Group of 178 Students in Grades Ten and Eleven.....	106
30. A Comparison of the Stanine Variations of Three or More in the Reading Performance of the Same Group of 200 Students As They Progressed From the Sixth Grade to the Eleventh.....	107
31. A Comparison of the Stanine Variations of Three or More in Arithmetic Reasoning Test Scores of the Same Group of 200 Students As They Progressed From the Sixth Grade to the Eleventh...	110
32. A Comparison of the Stanine Variations of Three or More in the English Scores of the Same Group of 200 Students As They Progressed From the Seventh Grade to the Eleventh.....	113
33. A Comparison of the Stanine Variations of Three or More in the Vocabulary Scores of the Same Group of 200 Students As They Progressed From the Seventh Grade to the Eleventh.....	116
34. A Comparison of the Stanine Variations of Three or More in the Arithmetic Computation and Spelling Scores of the Same Group of 200 Students in Grades Seven and Nine.....	117
35. A Comparison of the Stanine Variations of Three or More in the Achievement Scores of the Same Group of 200 Students When Different Grade Levels Are Used As Points of Reference.....	119

Table

Page

36. A Comparison of the Stanine Stability in the Reading Performance of the Same Group of 200 Students in Grades Six to Eleven When Different Grade Scores Are Used As Points of Reference.....120
37. A Comparison of the Stanine Stability in the Test Scores in Arithmetic Reasoning of the Same Group of 200 Students in Grades Six to Eleven When Different Grade Scores Are Used As Points of Reference.....121
38. A Comparison of the Stanine Stability in the English Scores of the Same Group of 200 Students in Grades Seven to Eleven When Different Grade Scores Are Used As Points of Reference.....122
39. A Comparison of the Stanine Stability in the Vocabulary Scores of the Same Group of 200 Students in Grades Seven to Eleven When Different Grade Scores Are Used As Points of Reference.....123
40. A Comparison of the Stanine Stability By Subject in the Performance of the Same Group of 200 Students in Grades Six to Eleven When All Possible Differentials Were Computed Between Subject Scores At Various Grade Levels.....124
41. A Comparison of the Stanine Stability in the Composite Scores of All Achievement Tests of the Same Group of 200 Students in Grades Six to Eleven When Different Tests Are Used As Points of Reference.....126
42. A Comparison of the Stanine Variations of Three or More in the Performance of the Same Group of 200 Students in Grades Seven to Eleven When Ability Is Used As a Reference Point and When Sixth-Grade Achievement Is Used As a Reference Point.....128
43. A Comparison of the Stanine Stability in the Composite Scores of the Same Group of 200 Students in Grades Six to Eleven When All Possible Differentials Were Computed Between All Scores at All Grade Levels.....131

Table	Page
44. Bivariate Distribution Showing Relationship Between Sixth-Grade Ability Stanines and High-School Average Achievement Stanines.....	135
45. Bivariate Distribution Showing Relationship Between Sixth-Grade Ability Stanines and Composite Stanine Scores on the Iowa Tests of Educational Development.....	136
46. Bivariate Distribution Showing Relationship Between Sixth-Grade Average Achievement Stanines and High-School Average Achievement Stanines...	137
47. Bivariate Distribution Showing Relationship Between Sixth-Grade Average Achievement Stanine Scores and ITED Composite Stanine Scores.....	139
48. A Comparison of the Stanine Variations in the Scores of the Same Group of 194 Students Between The California Test of Mental Maturity and All Achievement Tests in Grades Six, Seven, and Nine.....	161
49. A Comparison of the Stanine Variations in the Scores of the Same Group of 200 Students Between the Otis Test of Mental Ability and All Achievement Tests in Grades Nine, Ten, and Eleven.....	162
50. A Comparison of the Stanine Variations in All Ability-Achievement Scores of the Same Group of 200 Students As They Progressed From the Sixth Grade to the Eleventh.....	163
51. A Comparison of the Stanine Variations in the Reading Performance of the Same Group of 200 Students in Grades Six to Eleven When Different Grade Scores Are Used As Points of Reference...	164
52. A Comparison of the Stanine Variations in the Test Scores in Arithmetic Reasoning of the Same Group of 200 Students in Grades Six to Eleven When Different Grade Scores Are Used As Points of Reference.....	165
53. A Comparison of the Stanine Variations in the Composite Scores of Reading and Arithmetic Reasoning of the Same Group of 200 Students in Grades Six to Eleven When Different Grade Scores Are Used As Points of Reference.....	166

Table

Page

54. A Comparison of the Stanine Variations By Subject in the Performance of the Same Group of 200 Students in Grades Six to Eleven When All Possible Differentials Were Computed Between Subject Scores At Various Grade Levels..... 167
55. A Comparison of the Stanine Variations in the Composite Scores of All Achievement Tests of the Same Group of 200 Students in Grades Six to Eleven When Different Tests Are Used As Points of Reference..... 168
56. A Comparison of the Stanine Variations in the Composite Scores of the Same Group of 200 Students in Grades Six to Eleven When All Possible Differentials Were Computed Between Scores At All Grade Levels..... 169
57. A Summary of the Total Achievement Stanine Scores of the Same Group of 200 Students As They Progressed From Grade Six to Eleven..... 170

LIST OF FIGURES

Figure	Page
1. Stanine Chart.....	40
2. Stanine Chart Showing Identification of Students Whose Achievement Scores Deviated Significantly From Their Ability Scores.....	160

CHAPTER I

THE PROBLEM

1. Statement of the Problem

Measurement of students' ability and achievement through the use of standardized tests has now become an accepted part of nearly all school programs. However, in many cases the results of these tests are not put to optimum use because of the inability of the average teacher to interpret the comparative meaning of scores on different tests.

By the time a student reaches junior high school his permanent record card usually contains an accumulation of data such as measures of intelligence, achievement test scores, and teachers' marks, all recorded in various ways. Often a student's achievement is judged by simple inspection of the cumulative record. Thus, a student who scores somewhat above grade level may be considered to be a superior worker when actually he is achieving below his ability. If this underachievement happens to be in a major area such as reading, and goes unrecognized, the student may not accomplish his potential as he progresses through school because of the lack of basic skills.

As a junior high school counselor studies the elementary record cards, prior to the entry of the students, some of the questions and problems with which he is frequently confronted are:

1. Do these records indicate that the students are achieving within the range of their ability? How can this be determined?
2. Who are the superior students?
3. How can these superior students be challenged?
4. Will the students who are achieving at a high level in the sixth grade continue to do so as they progress through school?
5. Is the measure of intelligence or past achievement more important in the classification of students for instructional groups?

In an effort to find the answers to some of these questions the stanine stability of the test scores of 200 students was followed over a five-year period to determine how many of the students were accomplishing their potential as suggested by the sixth-grade records.

Purpose of the study.-- It was the purpose of this study to consider the use of national stanines as a methodology by which (1) any teacher with a limited knowledge of statistics could identify the potentially superior high school student before he reached junior

high school; (2) the students working above or below their ability could be easily recognized; (3) the records of students enrolled in the Providence Public Schools could be compared with those of students of comparable age and grade from other schools throughout the country.

Justification of the study.-- The potential of a superior student is often wasted because there is little challenge. Therefore, this type of student needs to be identified early and given an adequate school program in order to make effective use of his potential. Two important factors in the determination of a student's potential are mental ability and achievement as measured by high-quality standardized tests.^{1/} However, the administration of such tests is only one step in this process.

In the identification of potentially superior students it is necessary to know not only how the student compares with others in his own local school system, but also how he ranks with students of the same age and grade throughout the country. Frequently intelligence tests are interpreted by scores only, with little or no knowledge on the part of the interpreter of the method by which these scores were obtained, or of what they actually signify. An I. Q. of 100 on one standard test will generally mean the same thing as

1/John L. Holmes, Guidance Testing and the Identification of Pupil Characteristics, Educational Bulletin Number 21, California Test Bureau, Monterey, California, pp. 2-3.

100 on another test if similar standardization samples have been used. However, a score of 120 on one test may be at the 80th percentile and at the 90th percentile on another. It may also mean different things at different ages.^{1/} Teachers often interpret scores on achievement tests by grade equivalents only, completely overlooking the fact that a student may be one whole year ahead of his grade equivalent and yet be working below his ability.

These inconsistencies point up the need for a simple and reliable method of computing comparable measures by which the average teacher with little or no statistical training may evaluate the results of ability and achievement tests in terms of a student's potential.

Scope of the study.-- This study dealt with an analysis of repeated measurements of the same group of two hundred children during a span of five school years. Of this group one-half participated in a planned enriched program in junior high school, while the other half, who were of comparable age and ability, were given the regular course of study. Although it was not within the scope of this study to differentiate between the two groups, their

^{1/}H. S. Conrad, "Comparable Measures", pp. 279-282, Encyclopedia of Educational Research, W. S. Monroe, Editor, the Macmillan Company, New York, 1950.

identity was maintained throughout the study in order that the findings of this study might be used in a subsequent report on the comparative achievement of the two groups.

The California Short-Form Test of Mental Maturity^{1/} which was given in the first month of the sixth grade was the original criterion used in the selection of the students for this study. However, six students for whom this test score was not available were included in the study. These students did not enter the Providence schools until the seventh grade, at which time they became participants in the enriched program. The academic aptitude test used for these students was the Otis Quick-Scoring Mental Ability Test - Beta Form^{2/} which was given in the seventh grade. Of the entire group studied one hundred eighty-three students scored in the sixth stanine or higher on the original academic aptitude test.

Other tests used included the Otis Quick-Scoring Mental Ability Test^{3/} in the ninth grade, Metropolitan Achievement

1/E. T. Sullivan, W. W. Clark, and E. W. Tiegs, California Short-Form Test of Mental Maturity, Elementary, 1950-S Form, California Test Bureau, Los Angeles, California, 1950.

2/Arthur S. Otis, Otis Quick-Scoring Mental Ability Tests, Beta Form, World Book Company, New York, 1954.

3/Ibid.

Tests, ^{1/} Stanford Achievement Tests, ^{2/} Cooperative Tests, ^{3/} ⁶
and the Iowa Tests of Educational Development. ^{4/} The first
four represented all standardized tests taken by the entire
group of students in grades six, seven, nine, and ten. One
hundred twenty-one students took the Iowa Tests of Education-
al Development in grade eleven.

In order that the findings of this study might be of
practical value to the city of Providence, all tests used,
with the exception of the Iowa Tests of Educational Develop-
ment, were those which are administered at regular intervals
to all students in the public schools of Providence. The
Iowa Tests of Educational Development were included as
additional criteria of high-school achievement at the
eleventh-grade level.

1/Richard D. Allen, et al., Metropolitan Achievement Tests,
1947 Edition, Intermediate Battery-Partial, World Book
Company, New York.

2/Truman Kelley, et al., Stanford Achievement Test, Advanced
Battery-Partial, World Book Company, New York, 1953.

3/Educational Testing Service, Cooperative Tests, Princeton,
New Jersey, 1951.

4/E. F. Lindquist, The Iowa Tests of Educational Development,
Science Research Associates, Inc., Chicago, 1952.

2. Definitions of the Terms Used

The following are some of the terms used frequently in this study:

Academic aptitude.-- Combination of native abilities^{1/} and learned skills needed for success in school work.

Derived scores.-- Any set of transformed scores which get their meaning from the distribution characteristics of the test. Age equivalents, normalized standard scores, or percentile ranks are all derived scores.^{2/}

Deviation.--^{3/} The amount by which a test score differs from some stated reference value, e.g., deviations from the mean or deviations from the norm.

Stanine score.--^{4/} A simple nine-point scale of standard scores having a mean of five and a standard deviation of two.

Stanine differential.-- The difference between two stanine scores.

Stanine stability.-- A difference of two or less between two stanine scores.

^{1/}Walter N. Durost and William C. Kvaraceus, The Measurement Notebook, 1950, p. 1.

^{2/}Ibid., p. 2.

^{3/}Ibid., p. 3.

^{4/}Walter N. Durost, The Characteristics, Use, and Computation of Stanines, Test Service Notebook, World Book Company, New York, 1959, p. 1.

Significant deviation.-- A difference of three or more between two stanine scores.

Comparable scores.-- Scores from different tests which have been treated statistically so that they may be compared with one another.

Modal percentile norms.-- Percentile norms based upon a population of children who can be considered to have made normal progress through school, and who are not affected by acceleration and retardation practices.^{1/}

Overachievement.-- Score on achievement test three or more stanines above that on academic aptitude test.

Underachievement.-- Score on achievement test three or more stanines below that on academic aptitude test.

Plus achievement.-- Score on one achievement test three or more stanines above that on previous achievement test.

Minus achievement.-- Score on one achievement test three or more stanines below that on previous test.

Longitudinal study.-- A study of the same group of students through repeated measurements over a period of several years.

^{1/}World Book Company, Metropolitan Achievement Tests, Booklet of Norms, Section III, Grade and Percentile Norms, Modal Age Groups, Yonkers-on-Hudson, New York, 1947, p. 6.

3. Summary

This was a study to analyze the stability of national stanine scores on many different tests of both academic aptitude and achievement during a five-year period. The purposes were (1) to consider the reliability of using national stanine scores as comparable measures, (2) to evaluate the use of national stanine scores as possible criteria in the identification of the potentially superior high school student, and (3) to compare the records of students in the Providence schools with those of students of comparable age and grade throughout the country.

The measurements used consisted of scores on various standardized tests which were administered to the same group of two hundred students as they progressed from grade six to grade ten. One hundred twenty-one students were tested in the eleventh grade.

CHAPTER II
REVIEW OF RELATED LITERATURE

In his discussion on the broad purposes of educational and psychological testing Flanagan ^{1/} says in part:

"In a society such as ours where assisting the individual to attain the optimal development of his potentialities is one of the fundamental goals of the group, it becomes of extreme importance to supply the individual with all possible knowledge concerning himself, his knowledge, his skills, his abilities as compared with those of others, and his special areas of strength and weakness."

Another factor which might well have been included in the preceding paragraph under "all possible knowledge concerning himself" is his achievement as compared with his ability to achieve, for as Eckerson ^{2/} says, "Whether a pupil's achievement score is high or low has little significance without information about his school ability."

The task of supplying the student with knowledge of his achievement may be accomplished through the use of high-quality achievement tests which, according to Ebel, ^{3/} "...pro-

1/John C. Flanagan, "Units, Scores, and Norms," Chapter 17, Educational Measurement, E. F. Lindquist, Editor, American Council on Education, Washington, D. C., 1951, p. 761.

2/Louise O. Eckerson, "Testing and Counseling", School Life, (September, 1960), Volume 43:30.

3/Robert L. Ebel, Using Tests for Evaluation, Test Service Notebook, Number 15, Division of Test Research and Service, World Book Company, New York.

duce precise and relevant evidence on the degree of achievement of specified educational goals." However, the task of supplying knowledge of achievement in terms of the student's ability to achieve is a much more complex procedure. As Prescott ^{1/} states:

"The search for an acceptable method of comparing capacity to achieve with actual achievement has continued unrelentlessly from the time of the original accomplishment quotients of Franzen and Monroe and Buckingham. The fact that no less than nine different proposals applicable to the school grade level have been advanced is testimony to the need for such a technique and the complexity of the problem."

In his chapter on research Prescott ^{2/} describes and evaluates in detail the many capacity-achievement methods which have been proposed and used since the beginning of the testing movement. He concludes with the following statements:

"Many....revisions of the accomplishment quotient have been attempted....but because of the complexity of the whole problem, none has been entirely satisfactory. Indeed the problems in making such comparisons are so many and so serious that it is doubtful whether any method that may be devised can ever be completely satisfactory."

Before any method of comparing capacity with achievement can be considered it is first necessary to (1) understand the significance of the different types of scores and norms which are used in reporting results of educational and psychological tests, and (2) evaluate these scores and norms in terms of requirements for comparability.

1/George A. Prescott, The Development of An Improved Method of Making Capacity-Achievement Comparisons, Unpublished Doctoral Dissertation, Boston University, 1950, p. 38.

2/Ibid., p. 39.

For these reasons a brief review of the various ways in which test scores are expressed has been included in this chapter. A more detailed description of the stanine system has been given because of its importance in this study.

1. Ways of Expressing Test Scores

Test scores are meaningful and valuable to the extent that they can be interpreted in terms of capacities, abilities, and accomplishments of educational significance.^{1/}

Essentially, psychological test norms represent the test performance of the standardization sample, and provide basic reference points for the comparison of an individual's raw score on a test with that of the average performance of the standardization group.

In order to determine more precisely the individual's exact position with reference to the standardization sample, the raw score is expressed as a transformed score. These transformed scores serve two purposes: (1) they indicate an individual's relative standing in the normative sample and thus permit an evaluation of his performance in reference to other persons, and (2) they provide comparable measures which make possible a direct comparison of the individual's performance on different tests.

Since transformed scores can be expressed in the same units and referred to the same or to closely similar normative

^{1/}John C. Flanagan, op. cit., p. 695.

samples in the case of different tests, the individual's relative performance in many different functions can be compared.^{1/}

Each type of transformed or derived score has certain advantages and disadvantages, and in most situations the use of more than one type is necessary for adequate interpretation of test results.

Grade equivalents.-- One of the most widely used methods of interpreting test scores is that of the grade equivalent. The grade equivalent of any given test score is the grade level for which that score is the median score for all students at that level. Because of their apparent simplicity and ease of understanding, these scores are very popular with teachers, especially at the elementary grade levels. However, there are a number of reasons why they constitute a far-from-ideal type of interpretation at any grade level.^{2/} The authors of the Metropolitan Achievement Tests^{3/} point out some of these reasons in their manual for teachers:

"Traditionally, use of grade equivalents has been the most common method of interpreting results of elementary achievement tests. The simplicity of grade equivalents, their apparent comparability from subject to subject, and the continuity of the grade equivalent scale over the entire grade range commended this type of interpretation to many test users. There has been a growing awareness, however, that in some respects grade

^{1/}Anne Anastasi, Psychological Testing, The Macmillan Company, New York, 1955, pp. 72-73.

^{2/}John C. Flanagan, op. cit., p. 707.

^{3/}Walter M. Durost, (Editor), Directions for Administering, Metropolitan Achievement Tests, World Book Company, New York, 1959, p. 14.

equivalents are not entirely satisfactory values in which to express the achievement status of individual pupils in various areas, most notably in the case of pupils whose achievement differs markedly from the average or normal achievement. Units in the grade equivalent scale are not equal: six months of grade equivalent at one part of the scale may represent quite a different amount of achievement from six months at another part of the scale. Grade equivalents also are necessarily of unequal variability from subject to subject, which may lead to erroneous inferences concerning pupils' strengths and weaknesses. Experience has shown, too, that grade equivalents are subject to certain types of misinterpretation by parents and even by teachers with respect to appropriate pupil classification or grading. Despite these limitations grade equivalents have a very useful role to play in the interpretation of Metropolitan results, particularly in the interpretation of average results for classes, schools, or school systems. They facilitate the comparison of results for pupils in successive grades and indicate more directly than stanines or percentile ranks the amount of growth taking place in the various subjects."

Age equivalents.-- The age equivalent is similar to the grade equivalent, the difference being that age is substituted for grade as the basis for grouping the scores. Age norms are useful for the analysis of the performance of the individual in comparison with students of his own age regardless of grade placement. Their most extensive use is in the comparison of performance in a subject-matter area with the age values such as chronological age and mental age.^{1/}

Modal-age grade norms.-- Some of the disadvantages of age and grade equivalents are partially overcome by the use of another type of norm based on the scores of only those students in a grade who are typical with respect to age. This norm eliminates accelerated or retarded children who are below or

^{1/}John C. Flanagan, op. cit., p. 715.

above age for their grade, and has the effect of producing a more precisely defined group with which comparisons can be made. It also tends to set a higher standard of work.

Percentiles.-- Percentile scores are expressed in terms of persons in the standardization sample who fall below a given raw score. The percentile score provides a direct statement of the individual's relative position in the standardization sample. Percentile scores may be reported with reference to the total standardization sample, or they may be given separately for subgroups within the total sample. For example, percentile ranks are often reported in terms of age groups, grade groups, modal-age grade groups, and various other types.

Percentile scores are easy to compute and can be readily understood, even by relatively untrained people. They can be used equally well with children and adults, and are suitable for any type of test, whether it measures aptitude or personality variables.

The chief disadvantage of percentile scores is in the marked inequality of their units, especially at the extremes of the distributions. Under most conditions, percentile scores will give a correct picture of each individual's rank, or relative position, in the normative sample, but not of the amount of difference between his score and that of another person. For this reason, percentiles are not well suited to

computation of statistical measures.^{1/}

Standard scores.-- Another type of transformed score which is being used by current test makers is the standard score. This type of transformation will provide any desired mean and standard deviation for the derived scores by adding or subtracting a constant value to or from all raw scores and multiplying the result by another constant. In this type of transformation all differences between individuals retain their same relative values. Standard scores use a pair of constants which result in a mean of zero and a standard deviation of one for the group used as a standard. Variations of this technique use mean values such as 50, 100, or 500, with standard deviations of 10, 20, and 100 respectively. These types of scores simplify interpretation and increase comparability. However, standard scores of this type are comparable only when found from distributions which have approximately the same form. If the raw-score scale is distorted, the standard scores will have the same defect.^{2/}

In order to achieve comparability of scores from dissimilarly shaped curves, nonlinear transformations may be used to fit the scores to any specified type of distribution curve. Normalized standard scores are standard scores in a distribution

^{1/}Anne Anastasi, op. cit., pp. 77-79.

^{2/}John C. Flanagan, op. cit., pp. 722-723.

which has been transformed to fit a normal curve. Like linear standard scores, these scores have a mean of zero and a standard deviation of one. However, since the distribution has been transformed to fit a normal curve, a normalized standard score of zero indicates that the student falls at the mean of a normal curve, excelling 50 per cent of the group.

As with linear transformations, normalized standard scores can be put into any convenient form.^{1/}

Stanine scores.-- One of the most recent and for many purposes the most useful types of transformation is the stanine scale. This method of interpreting test scores was first used by the Air Force during World War II when it was necessary to translate mass test data into a simple workable form. Flanagan^{2/} describes the procedure as follows:

"The procedure selected for use in the Aviation Psychology Program was the assignment of weights to the various test scores and the computation of a weighted average score for each individual. Since the importance of the various tests was generally different for each of the various air-crew positions, there were as many weighted averages obtained as there were different air-crew positions for which the man might be classified.

^{1/}Anne Anastasi, op. cit., pp. 81-83.

^{2/}John C. Flanagan, Editor, The Aviation Psychology Program in the Army Air Forces, Report Number 1, Army Air Forces Aviation Psychology Program Research Reports, U. S. Government Printing Office, Washington, D. C., 1948, pp. 66-67 (As cited by John E. Corcoran, A Longitudinal Study to Determine the Stanine Stability of a Group's Test-Score Performance in the Elementary School, Unpublished Doctoral Dissertation, Boston University School of Education, 1958, p. 22.)

To simplify the interpretation and use of these weighted average aptitude scores, they were converted to a simple and convenient nine-point scale. This was done by assigning various percentages of the total distribution to the various score intervals beginning with the lowest 4 per cent receiving scores of 1 and going on until the highest per cent received scores of 9.

These values represented normalized standard scores with a mean of five and a standard deviation of approximately two on a nine-point scale. The term stanine was coined to facilitate reference to these scores."

The stanine method retains many of the advantages and eliminates many of the disadvantages of other types of scores. The single-digit scores provide a maximum in the way of simplicity and ease of handling and interpretation. Each unit is of equal value and immediately identifies the relative standing of a student within a group. Scores between different tests may be directly compared provided the standardization sample is the same or closely similar. Through the use of subgroup norms stanine scores make it possible to obtain the relative standing of an individual with respect to all types of groups such as his standing in his class, his school, and nationally. The stanine system also permits evaluation of achievement in relation to ability or previous achievement.

One of the disadvantages of the stanine system is that the 1 per cent of the students at either end of the curve are grouped together. As Guilford ^{1/} points out, in guidance many counselors would not want the rare person at either end of

1/J. P. Guilford, Fundamental Statistics in Psychology and Education, Second Edition, McGraw-Hill Company, New York, 1950, p. 305.

the curve submerged with the other 3 per cent, as there is probably a full unit's discrimination between the end 1 per cent and the next 3 per cent. Flanagan^{1/} describes stanine scores as, "....in general....too coarse to preserve all of the information contained in the raw scores on educational achievement tests."

Some of the disadvantages of stanine scores may be minimized in the following ways:

1. Both percentile ranks and stanine scores may be recorded as was done in this study.
2. An 11-point scale which separates stanines 1 and 9 into two equal parts of 1 per cent and 3 per cent respectively may be used.
3. Where accuracy of measurement justifies greater refinement of units for a specific purpose, each unit may be divided into three equal parts.^{2/}

2. Stanine Studies

Although no study on the use of national stanines could be located in the research, several studies were found which involved the use of local stanines.

Corcoran's^{3/} study, which was the only longitudinal one

^{1/}"Units, Scores, and Norms", op.cit., p. 747.

^{2/}Ibid.

^{3/}John E. Corcoran, A Longitudinal Study to Determine the Stanine Stability of a Group's Test-Score Performance in the Elementary School, Unpublished Doctoral Dissertation, Boston University School of Education, 1958.

found in the literature, was an analysis of the ability and achievement test results of the same 336 elementary school children over a seven-year period. He converted all test scores of students in one school system into local stanines through the use of Otis Normal Percentile Charts and then studied the seven-year stanine stability of the 336 students who constituted the "staying population". His data included the test scores on three different ability tests, Pintner General Ability Test, Kuhlmann-Anderson Intelligence Tests, and the SRA Primary Mental Abilities Test; and the yearly scores on the sub-tests of the Metropolitan Achievement Tests-Partial Battery from grades two through eight inclusive. His purposes were to determine the stanine stability of the group's performance in ability, achievement as related to ability, and achievement as related to previous achievement.

His findings indicated that the peak of mental stability was reached somewhere at or near the fifth grade, and that there was a substantial degree of stability at all levels tested. When ability was compared with achievement, vocabulary showed the most even pattern at all grade levels. Reading showed the highest stability (88 per cent) in grade five, with stabilities of 80 and 81 per cent at grades two and eight respectively. The significant deviations in reading were more positive than negative. The lowest average stability, 74 per cent, was found in grade two, and the high-

est, 85 per cent, at grade five.

When achievement was compared with previous achievement a grade to grade stability of 90 per cent or higher was found in reading, vocabulary, English, and spelling. Arithmetic fundamentals and problems gradually increased from 77 and 76 per cent at grade three to 85 and 91 per cent at grade five. From grade five on all subjects reached a 90 per cent stability except arithmetic fundamentals and English at grade seven. The greatest percentage of minus achievement was found at the seventh-grade level.

The results of Corcoran's study indicated that the majority of students in a staying population will show stanine stability in (1) ability, (2) achievement as related to ability, and (3) achievement as related to previous achievement during the elementary school years. His results also indicated that scores on previous achievement tests are better predictors of future achievement than those on mental ability tests.

Another study which was based on the use of local stanines is that of Brown.^{1/} This was an analysis of the ability and achievement test results of 121 fifth-graders in a small town in Massachusetts. The purpose was to determine the number of

1/Winnifredian Brown, A Stanine Method of Making Ability-Achievement Comparisons, Unpublished Master's Thesis, Boston University School of Education, 1958.

pupils who were achieving above, below, or at maximum capacity according to the stanine scale. The ability test used was the Kuhlmann-Anderson Intelligence Tests, and the achievement measures were the sub-tests of the Stanford Achievement Test-Partial Battery. All ability and achievement test scores were converted into stanines through the use of Otis Normal Percentile Charts. The ability stanine was then compared with the achievement stanine with the following results:

1. Reading and word meaning showed the greatest consistency with stabilities of 94.2 and 90.9 per cent respectively.
2. Arithmetic computation indicated the lowest stability, 79.3 per cent, with more pupils scoring above and below ability than in any other area.

In conclusion Brown states that long term records indicate adherence to a given stanine for both intelligence and achievement.

3. Reports on Applications of Stanines

Within the past few years several school systems have been experimenting with the various ways in which stanines may be used. Some of these reports which are not particularly applicable to this study but may prove useful to the reader who is interested in further information regarding specific

applications of stanines are those by Hart,^{1/} Durost,^{2/} and Sexton.^{3/}

During the progress of this study a new edition of the Metropolitan Achievement Tests was published which, for the first time, included stanines as a method of interpreting test scores. The following evaluation of the stanine method is found in their guide for examining the new test series:^{4/}

"This system has been selected as the simplest, most readily understood, and most functional method of score interpretation yet devised. In the stanine system, scores are converted to a simple nine-point equal-step scale.

New and more rigorous comparability is now possible: the child is compared with other children in his own group; scores are compared from one test to another; and the child's growth is measured from year to year.

The Metropolitan stanine scale allows identification and grouping in equal ranges of ability. It permits more ready and more dependable identification of talented, average, and slow-learning groups. It simplifies any statistical treatment and analysis of test data, such as relating one test to another, or to other variables such as teachers' marks.

Metropolitan stanine scores may be readily incorporated into the teacher's mark or grade. Particularly valuable is the comparison they permit between achievement and capacity. They provide a graphic, easily

1/Irene Hart, Using Stanines to Obtain Composite Scores Based On Test Data and Teachers' Ranks, Test Service Bulletin, No. 86, Division of Test Research and Service, World Book Company, New York, 1957.

2/Walter N. Durost, The Characteristics, Use, and Computation of Stanines, Test Service Notebook, Number 23, Division of Test Research and Service, World Book Company, New York, 1959.

3/John M. Sexton, Analyzing Student Achievement in Tenth Grade English, Test Service Bulletin, No. 92, Division of Test Research and Service, World Book Company, New York, 1959.

4/Harold H. Bixler et al., Metropolitan Achievement Tests, World Book Company, New York.

understood record of the child's performance in the various subtests, which, at the school's discretion, may be presented to parents. MAT stanine scores have been well received by parent groups, because they show where a child stands in relation to other children in his class, illuminate his strengths and weaknesses, and thus broaden the usefulness of test information."

As pointed out in the above evaluation, there are many advantages to be derived in the use of the stanine method of reporting test scores. If the practical application of this new technique proves as effective as the test authors predict the next decade may find many more test makers including stanines as a form of interpretative score.

CHAPTER III

PROCEDURES

1. Student Population

This study was a follow-up of a group of 200 students in the Providence Public School System whose sixth-grade ability or achievement records were suggestive of future academic success. These students entered junior high school in September of 1954, at which time they were placed in homogeneous instructional groups according to the following criteria: (1) the California Short-Form Test of Mental Maturity;^{1/} (2) the Metropolitan Achievement Tests^{2/} of reading and arithmetic problems; and (3) the teachers' ratings in reading and arithmetic. The standardized tests were given in October of 1953, and the teachers' ratings in January of 1954. At this time all students were in the first half of the sixth grade.

The first criterion used for grouping was the I.Q. score obtained on the California Short-Form Test of Mental Maturity. The I.Q. is considered the most popularly known intelligence concept. It is a measure of a person's potential

^{1/}E. T. Sullivan, op. cit.

^{2/}Richard D. Allen, op. cit.

and an indicator of what his maximum ability level will be at full maturity. It is a ratio which indicates the rate at which an individual is developing mental ability.^{1/}

Since, normally, the level at which an individual is able to score on an intelligence test (in relation to scores of others of comparable age) remains fairly constant throughout his life, the intelligence quotient "can be used to predict future mental age or ability at any time, up to and including adulthood."^{2/}

Ability to score on an intelligence test is related to success in school. However, achievement in school, although closely related to general intelligence, is by no means identical with it, "but results from a combination of other factors such as general maturity, motivation, diligence, and efficiency of study habits."^{3/}

The other criteria used were the sub-tests of reading and arithmetic problems of the Metropolitan Achievement Tests, and the teachers' rating in reading and arithmetic. All criteria were based on 6B records.

1/Elizabeth T. Sullivan, Willis W. Clark, and Ernest W. Tiegs, California Short-Form Test of Mental Maturity, Elementary Grades, Manual for Interpreting, California Test Bureau, Los Angeles, 1957, p. 14.

2/Ibid.

3/N. E. A. Educational Policies Commission, Education of the Gifted, National Education Association, Washington, D. C., 1950, pp. 35-46.

At the end of two years 76 students enrolled in Classical High School for ninth-grade work, while the other 124 remained in junior high school for one more year, and then enrolled in Hope High School, a comprehensive high school.

2. Methodology

Because of the large number of scores used in the study, it was decided to convert all scores into stanines. No reference to national stanine conversion was located in the research, as all previous studies on stanine involved the use of local percentiles. However, because the majority of this group were above average in both ability and achievement, it was considered important to know how they ranked with students of comparable age and grade throughout the country. For this reason national modal percentile norms were converted into national stanines. In spite of the fact that the standardization groups of the various tests may not have been identical, it was assumed that the Providence students were sufficiently cosmopolitan to allow for varying samples.^{1/}

3. Source of the Data

The data used in this study were procured from the

1/An additional factor was the participation of local schools in the standardization of several of the tests used.

results of intelligence and achievement tests given in the Providence schools during a six-year period. The series of tests began when the students were in the second month of the sixth grade, and included the successive years through the second month of the junior year in high school.

Intelligence measures.-- Two measures of intelligence or academic aptitude were given during the span of this study:

1. Grade Six, October, 1953, California Short-Form Test of Mental Maturity, Elementary Grades 4-8, 1950 S-Form ^{1/}
2. Grade Nine, October, 1956, Otis Quick-Scoring Mental Ability Test, Beta Test-Form E ^{2/}

Achievement measures.-- The achievement measures used in grades six to nine were The Metropolitan Achievement Tests ^{3/} and The Stanford Achievement Tests. ^{4/} The tests given, date of administration, and grade were as follows:

1. Grade Six, October, 1953, Metropolitan Achievement Tests, Intermediate Battery-Partial Form
2. Grade Seven, October, 1954, Stanford Achievement

^{1/}E. T. Sullivan, op. cit.

^{2/}Arthur S. Otis, op. cit.

^{3/}Richard D. Allen, op. cit.

^{4/}Truman Kelley, op. cit.

Tests, Advanced Battery-Partial Form L

3. Grade Nine, October, 1956, Stanford Achievement

Tests, Advanced Battery-Partial Form J

From the second half of the ninth grade through the tenth, the measurements of achievement consisted of a series of Cooperative Tests:^{1/}

1. Grade Nine, May, 1957

Elementary Latin Test, Form Q

Elementary French Test, Form Q

Elementary Algebra Test, Form Z

2. Grade Ten, October, 1957

English Test A1, Mechanics of Expression,
Form Z

English Test C1, Reading Comprehension,
Form T

3. Grade Ten, May, 1958

English Test B1, Effectiveness of Expression,
Form T

Algebra, Intermediate, Form Z

French, Elementary, Form R

Latin, Advanced, Form R

Biology, Form Y

Modern European History, Form X.

1/op. cit.

The Iowa Tests of Educational Development ^{1/} were used as achievement measures in the eleventh grade. These were administered in October, 1958.

4. Description and Evaluation of the Measuring Instruments Used

In discussing capacity-achievement comparisons
Prescott ^{2/} says in part:

"Any discussion of the difficulties involved in making capacity-achievement comparisons must commence by considering the validity of the capacity and achievement measures being compared, for no method, regardless of how elaborate it may be, can compensate for lack of validity in the measures employed."

California Short-Form Test of Mental Maturity.-- The California Short-Form Test of Mental Maturity is part of the larger test called the California Test of Mental Maturity. It is available at five levels from pre-primary to advanced.

This short form of the test has been developed to secure as valid a measure of mental maturity as can be secured by a one-period group test.

Selections from the longer form of the test provide sub-tests which measure both language and non-language mental maturity, and four of the major factors involved in intelligence or mental capacity, namely: spatial relations,

1/E. F. Lindquist, op. cit.

2/George A. Prescott, The Development of an Improved Method of Making Capacity-Achievement Comparisons, Unpublished Doctor's Dissertation, Boston University, 1950, p. 40.

logical reasoning, numerical reasoning, and verbal concepts which are useful in the thinking process.

Because of the wide range of abilities found in most age or grade groups this test provides for measurements several grades or years above and below the particular group being tested.^{1/}

The validity of these tests was established through intercorrelations with other intelligence tests, correlations with achievement tests, intercorrelations among parts of the test itself, and factor analytic evaluation.^{2/}

The total mental factors score has been found to correlate as high or higher with the individual Stanford-Binet than any other one-period mental ability test.^{3/}

The norms for the California Short-Form Test of Mental Maturity were based on a controlled sampling of over 125,000 cases from representative school districts throughout the United States, which constituted a normal distribution of mental ability, typical age-grade relationships, and 15 per cent minority groups.^{4/}

^{1/}Elizabeth T. Sullivan, Willis W. Clark, and Ernest W. Tiegs, California Short-Form Test of Mental Maturity, Elementary, 1950 S-Form, Manual, California Test Bureau, Los Angeles, California, 1950, p. 2.

^{2/}James C. Coleman, Summary of Investigations Number Three, California Test of Mental Maturity, California Test Bureau, Los Angeles, California, 1956, p. 6.

^{3/}Elizabeth T. Sullivan, op. cit., p. 4.

^{4/}Ibid., p. 15.

Otis Quick-Scoring Mental Ability Test.-- The purpose of the Otis Quick-Scoring Mental Ability Test is to measure mental ability-thinking power or the degree of maturity of the mind.^{1/}

The determination of the validity of each item in the test consisted of comparing the number of passes of that item by a group of pupils who were making rapid progress through school with the number of passes of the item by a group of students who were making slow progress through school. Only those items were used which showed a distinct gain in the number of passes of the rapid-progress students over the number of passes of the slow-progress students. Each item justified its inclusion because it contributed definitely to the capacity of the test to measure brightness as reflected in the rate of progress through school.^{2/} The correlation between the Otis scores and scores on the sub-tests of the Stanford Achievement Test ranged from .564 to .827.

Metropolitan Achievement Tests.-- Hildreth^{3/} describes these tests in the following manner:

^{1/}Arthur S. Otis, Otis Quick-Scoring Mental Ability Tests, Manual of Directions for Beta Test, World Book Company, New York, 1954, p. 1.

^{2/}Ibid., p. 8.

^{3/}Gertrude Hildreth, Manual for Interpreting Metropolitan Achievement Tests, World Book Company, New York, 1948, p. 5.

"The Metropolitan Achievement Tests series is a comprehensive series of educational achievement tests consisting of five batteries covering the essential skill subjects and content areas taught in Grades 1 to 9. Each battery consists of a group of different tests which were subjected to the same experimental verification and were standardized at the same time, on the same pupil population. The several forms of each battery are similar in content and organization and are equal in difficulty."

The tests were restandardized in 1947 in a nation-wide program which involved the administration of 500,000 tests to students in a variety of communities in every state in the country. The actual norms were based on a 25 per cent random sample from each classroom tested. The normalized scaled scores were based on the distribution of raw scores for the sixth-grade modal age group.

In discussing the validity of these tests Hildreth states ^{1/} that no material was included which could not be justified in terms of representative courses of study, popular textbooks, and opinions of experts in the various fields. These series of tests are considered to be valid of typical content of instruction in this country.

Although a partial battery of these tests is given each year, the results of only two of the sub-tests were used in this study. These were reading and arithmetic problems, and were taken from the scores obtained in the sixth grade.

1/op. cit., p. 8.

Stanford Achievement Tests.--- The Stanford Achievement Tests are a series of comprehensive achievement tests designed to measure the important knowledges, skills, and understandings commonly accepted as desirable outcomes of the major branches of the elementary curriculum.

They are available in five equivalent forms which have been matched for content and difficulty, represent equally good measures of the respective subjects, and yield directly comparable results.^{1/}

The content of these tests is based on detailed analysis of contemporary textbooks, courses of study, and professional literature in each of the various subject fields.^{2/}

The Stanford Achievement Tests were standardized in 1952 in a national program which included the testing of 460,000 pupils in 363 school systems drawn from thirty-eight states.^{3/} The final norm group included 25 per cent random sample of 340 school systems.

In this study the Advanced Battery-Partial was used which included tests in the following areas:

Paragraph Meaning	Language
Word Meaning	Arithmetic Reasoning
Spelling	Arithmetic Computation

1/Truman Kelley, et al., Stanford Achievement Test, Directions for Administering, World Book Company, New York, 1953, pp. 1-2.

2/op. cit.

3/Ibid., p. 15.

The results of all tests for Grades 7, 8, and 9 were recorded originally, but it was later decided to omit the eighth grade scores.

Cooperative Achievement Tests.-- The Cooperative Achievement Tests include a wide variety of selected subject-matter tests in fields such as English, foreign languages, mathematics, science, and social studies. Tests appropriate for high-school and college students are available. The reliability is reported in terms of standard error of measurement for each test, and coefficients are usually above 0.90. Test items are based on analysis of textbooks and are reviewed by subject-matter specialists.^{1/}

In preparation of the norms for secondary schools, each school system was classified according to geographical location, and size of community. Preliminary investigation indicated that the performances of the groups in the East, Middle West, and West were sufficiently similar to justify grouping them under one heading.^{2/}

Percentile ranks for each grade level in three types of secondary schools and for three types of colleges are provided for the English tests. Percentile ranks for each grade level are available in other subjects.

^{1/}Clifford P. Froehlich and John G. Darley, Studying Students, Science Research Associates, Inc., Chicago, 1952, p. 255.

^{2/}Educational Testing Service, Introduction to the Norms, Cooperative Achievement Tests, Princeton, New Jersey, p. 3.

In his review of The Cooperative English Test, Reading Comprehension, J. B. Stroud says: "It would be difficult to find a test better conceived and better executed within the limits of its objectives."^{1/}

Robert C. Pooley considers the Cooperative English Test one of the best tests available in the field of English skills.^{2/}

Iowa Tests of Educational Development.-- The Iowa Tests of Educational Development are a battery of nine objective tests which provide meaningful data on the individual capabilities of students. Over 700 items are included, each selected for its discriminating power in evaluating the lasting results of an educational program. It is not primarily a test of subject matter. Emphasis is placed on broad intellectual skills and interests, on understanding of and ability to use what is learned, rather than on sheer knowledge of facts.^{3/}

The individual tests in the battery are as follows:^{4/}

^{1/}Oscar K. Buros (Editor), The Third Mental Measurements Yearbook, Rutgers University Press, New Brunswick, New Jersey, 1949, p. 498.

^{2/}Ibid., p. 120.

^{3/}Science Research Associates, Inc., Catalog, 259 E. Erie Street, Chicago 11, Illinois, 1960, p. 11.

^{4/}Ibid.

1. Understanding of Basic Social Concepts
2. Background in the Natural Sciences
3. Correctness and Appropriateness of Expression
4. Ability to do Quantitative Thinking
5. Ability to Interpret Reading Materials in the Social Studies
6. Ability to Interpret Reading Materials in the Natural Sciences
7. Ability to Interpret Literary Materials
8. General Vocabulary
9. Uses of Source Information

The composite score gives a fairly accurate picture of the student's general level of educational development.

The norms of the Iowa Tests of Educational Development are based on a nationwide testing program conducted in 1957.^{1/} A total of 366 schools in 254 school systems participated in the program. These systems were drawn from all 48 states, and included 148,590 pupils in grades nine through twelve.

"In constructing these tests, the authors prepared a considerably larger number of items than was required for the finished forms. All of this material was tried out in advance in a representative sample of several hundred Iowa high schools. The items for the finished tests were then selected for discriminating power and appropriate level of difficulty on the basis of a statistical analysis of the tryout data."^{2/}

Each of the Iowa Tests of Educational Development was designed to yield a coefficient of reliability of 0.91 for students in a single grade in a single school.

^{1/}Science Research Associates, Inc., Iowa Tests of Educational Development, Manual for the School Administrator, Chicago, 1957, pp. 32-36.

^{2/}Manual for School Administrator, op. cit., p. 28.

The Iowa Tests were given to 121 selected students in the eleventh grade.

4. Summary of the Measuring Instruments

So far a brief description and evaluation of the measuring instruments used in this study have been presented. Both the ability and achievement measures have been shown to be valid, reliable, and applicable to the Providence schools.

The standard or scaled scores of all these tests provide the basic framework for this study. The total number of pupil test scores, including sub-test scores, used during this study was 5461.

5. Organization and Treatment of the Data

Recording of the data.-- All data were recorded on 4" by 6" cards. Four different colors were used to designate the four different groups--Classical criterion, Classical matched, Hope criterion, Hope matched.^{1/} Through the use of colors the cards could be combined in any manner and still be easily identified by groups.

1/The identity of the groups was maintained for a subsequent report. The criterion group consisted of 100 students from one junior high school who participated in a planned enriched program. The matched group consisted of 100 students of comparable age and ability from all other junior high schools in the city.

For each academic aptitude test the following data were recorded: raw score, chronological age, mental age, I.Q. score, national modal percentile rank, and national stanine score. The data used for achievement tests included: grade or standard score, grade equivalent, national modal percentile rank, and national stanine score. Percentile ranks were taken from the tables of national modal percentile norms for each test. The percentile ranks were then converted into national stanine scores through the use of the stanine scale as shown in Table 1.

Table 1. Proportion of Cases, and National Modal Percentile Ranks Included in Each National Stanine Level

National Stanine level	Proportion of cases	National modal percentile ranks
(1)	(2)	(3)
9.....	4 per cent	96 to 99
8.....	7 per cent	89 to 96
7.....	12 per cent	77 to 89
6.....	17 per cent	60 to 77
5.....	20 per cent	40 to 60
4.....	17 per cent	23 to 40
3.....	12 per cent	11 to 23
2.....	7 per cent	4 to 11
1.....	4 per cent	1 to 4

Plotting of stanine charts.-- After all data had been recorded and all scores converted to national stanine ranks, the next step was to compare the stanine scores obtained on the different tests by each individual student. This was done through the use of a stanine chart which is shown in Figure 1.

The double lines in Figure 1 represent the same stanine for both ability and achievement, as shown by numbers 1, 2, and 3. The dotted lines indicate variations of two or less between ability and achievement. The squares on the left show plus variation; and those on the right, minus. Numbers 4 and 5 indicate plus variations, and numbers 6 and 7 indicate minus variations. All scores which fall outside the dotted lines show significant deviations. The scores to the left of the dotted lines indicate overachievement, as shown by numbers 8 and 9. The scores to the right of the dotted lines indicate underachievement, as shown by numbers 10 and 11.

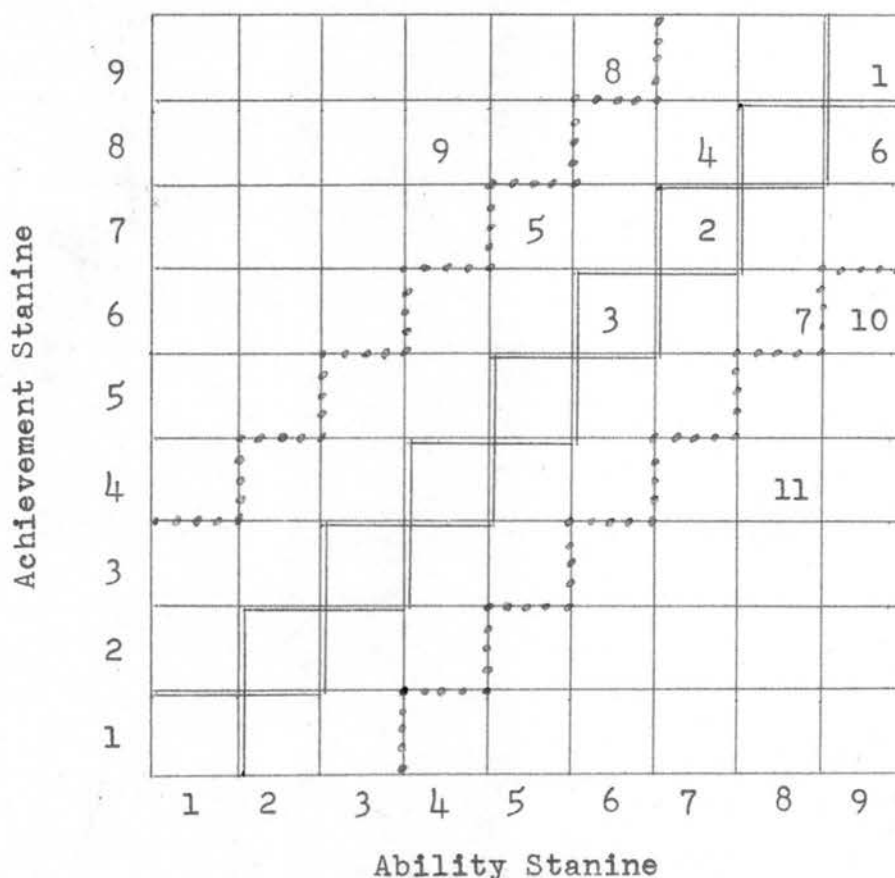


Figure 1. Stanine Chart

Ability versus achievement.-- The stanine scores obtained on the California Short-Form Test of Mental Maturity when the students were in the sixth grade were projected through the eleventh grade, and stanine differentials were obtained between the following tests as seen in Table 2:

1. Metropolitan Achievement Tests - Grade Six

Reading
Arithmetic Reasoning

2. Stanford Achievement Tests - Grades Seven and Nine

Paragraph Meaning
Arithmetic Reasoning
Word Meaning
Language Usage
Arithmetic Computation
Spelling

3. Iowa Tests of Educational Development

Composite Score.

The stanine scores obtained in the ninth grade on the Otis Quick-Scoring Test of Mental Ability were projected through the eleventh grade, and differentials computed between the following tests and sub-tests:

1. Stanford Achievement Tests - Grade Nine

Paragraph Meaning
Arithmetic Reasoning
Word Meaning
Language Usage
Arithmetic Computation
Spelling

2. Cooperative Tests

Elementary Algebra
Intermediate Algebra
Elementary Latin

Table 2. Tests Used in Computation of Ability-Achievement
Stanine Differentials

Achievement Tests and Sub-tests (1)	Points of Reference	
	California (2)	Otis (3)
Metropolitan-Grade 6		
Reading Comprehension.....	x	
Arithmetic Problems.....	x	
Stanford-Grade 7		
Paragraph Meaning.....	x	
Arithmetic Reasoning.....	x	
Language Usage.....	x	
Word Meaning.....	x	
Arithmetic Computation.....	x	
Spelling.....	x	
Stanford-Grade 9		
Paragraph Meaning.....	x	x
Arithmetic Reasoning.....	x	x
Language Usage.....	x	x
Word Meaning.....	x	x
Arithmetic Computation.....	x	x
Spelling.....	x	x
Cooperative-Grades 9A, 10		
Elementary Algebra.....		x
Elementary Latin.....		x
Elementary French.....		x
Intermediate Algebra.....		x
Advanced Latin.....		x
Reading Comprehension.....		x
Mechanics of Expression.....		x
Effectiveness of Expression.....		x
Biology.....		x
Modern European History.....		x
Iowa Tests of Educational Development-Grade 11		
Basic Social Concepts.....		x
Background in Natural Sciences.....		x
Correctness and Appropriateness of Expression.....		x
Quantitative Thinking.....		x
Reading in Natural Sciences.....		x
Reading in Social Studies.....		x
Literary Materials.....		x
General Vocabulary.....		x
Source Information.....		x
Composite Score.....	x	x

Advanced Latin
 Elementary French
 Reading Comprehension
 Mechanics of Expression
 Effectiveness of Expression
 Biology
 Modern European History

3. Iowa Tests of Educational Development - Grade Eleven

Understanding of Basic Social Concepts
 Background in Natural Sciences
 Correctness and Appropriateness of Expression
 Ability to do Quantitative Thinking
 Ability to Interpret Reading Materials in the
 Natural Sciences
 Ability to Interpret Reading Materials in the
 Social Studies
 Ability to Interpret Literary Materials
 General Vocabulary
 Uses of Source Information
 Composite Score.

Achievement versus achievement.-- Stanine differentials
 in achievement from grade to grade were computed between
 the following tests and sub-tests which are listed in Table
 3:

1. Metropolitan Achievement Tests - Grade Six

Reading

Stanford Achievement Test Paragraph Meaning
 Grade Seven
 Stanford Achievement - Paragraph Meaning
 Grade Nine
 Cooperative Reading Comprehension - Grade 10B
 Iowa Tests of Educational Development - Ability
 to Interpret Literary Materials - Grade 11B

Arithmetic Problems

Stanford Achievement - Arithmetic Reasoning
 Grade Seven B

Table 3. Tests Used in the Computation of Achievement-Achievement Stanine Differentials

Achievement Tests and Sub-Tests	Points of Reference															
	RC	AP	PM	AR	E	WM	AC	SP	PM	AR	E	WM	MA	RC	EM	EE
	6	6	7	7	7	7	7	7	9	9	9	9	9A	10	10	10
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
Stanford-Grade 7																
Paragraph Meaning.....	X															
Arithmetic Reasoning.....		X														
Stanford-Grade 9																
Paragraph Meaning.....	X		X													
Arithmetic Reasoning.....		X		X												
Language Usage.....					X											
Word Meaning.....						X										
Arithmetic Computation.....							X									
Spelling.....								X								
Cooperatives-Grades 9A, 10B																
Elementary Algebra.....		X		X						X						
Intermediate Algebra.....										X			X			
Reading Comprehension.....	X		X						X							
Mech. of Expression.....					X						X					
Effect. of Expression.....					X						X				X	
Iowa Tests of Educational Development-Grade 11																
Test 3.....					X						X				X	X
Test 4.....		X		X						X			X			
Test 7.....	X		X						X					X		
Test 8.....						X						X				

Stanford Achievement - Arithmetic Reasoning - Grade 9B
 Cooperative Elementary Algebra - Grade 9A
 Iowa Tests of Educational Development - Ability to do Quantitative Thinking - Grade 11B

2. Stanford Achievement Tests - Grade Seven

Paragraph Meaning

Paragraph Meaning - Grade 9B
 Cooperative Reading Comprehension - Grade 10B
 Iowa Tests of Educational Development - Ability to Interpret Literary Materials - Grade 11B

Arithmetic Reasoning

Arithmetic Reasoning - Grade 9B
 Cooperative Elementary Algebra - Grade 9A
 Iowa Tests of Educational Development - Ability to do Quantitative Thinking - Grade 11B

Word Meaning

Word Meaning - Grade 9B
 Iowa Tests of Educational Development - Vocabulary - Grade 11B

Language Usage

Language Usage - Grade 9B
 Cooperative - Mechanics of Expression - Grade 10B
 Cooperative - Effectiveness of Expression - Grade 10A
 Iowa Tests of Educational Development - Correctness and Appropriateness of Expression - Grade 11B

Arithmetic Computation and Spelling

Arithmetic Computation and Spelling - Grade 9B

3. Stanford Achievement Tests - Grade Nine B

Paragraph Meaning

Cooperative - Reading Comprehension - Grade 10B
 Iowa Tests of Educational Development - Ability

to Interpret Literary Materials - Grade 11B

Arithmetic Reasoning

Cooperative - Elementary Algebra - Grade 9A
Iowa Tests of Educational Development - Ability
to do Quantitative Thinking - Grade 11B

Language Usage

Cooperative - Mechanics of Expression - Grade
10B
Cooperative - Effectiveness of Expression -
Grade 10A
Iowa Tests of Educational Development -
Correctness and Appropriateness of Expression -
Grade 11B

Word Meaning

Iowa Tests of Educational Development -
Vocabulary

4. Cooperative Tests

Mechanics of Expression - Grade Ten B

Effectiveness of Expression - Grade 10A
Iowa Tests of Educational Development -
Correctness and Appropriateness of Expression -
Grade 11B

Effectiveness of Expression - Grade Ten A

Iowa Tests of Educational Development -
Correctness and Appropriateness of Expression -
Grade 11B.

Ability versus ability.-- Stanine differentials in

ability were computed between the following tests:

1. California Short-Form Test of Mental Maturity -
Grade Six B
2. Otis Quick-Scoring Test of Mental Ability - Grade 9B.

In order to maintain the identity of the four separate

741-22

groups, different colored pencils were used in the plotting of the stanine charts. Whenever a differential of three or more existed between stanine scores the student's number was placed beside the tallies in the square.^{1/} This method made it possible to classify the results in three different ways:

1. The number of students who were achieving above or below their ability or previous achievement in the various tests could be identified.
2. The number of different students who were involved in the significant deviations could be computed.
3. The schools from which these students came could be identified.

To simplify these procedures, group check charts were used. A blue check mark signified a plus deviation, and a red one a minus deviation. The degree of deviation was also recorded above the check mark. In this manner the entire record of stanine deviations of both the individual student and the group could be seen at a glance.

6. Summary of Procedures

This was a follow-up study of the records of a group of 200 students as they progressed from the sixth grade to the junior year in high school.

1/See appendix.

III. 23

The purposes of the study were:

1. To determine how many students were achieving according to their ability as measured in the sixth grade
2. To determine how many students remained at a comparable level of achievement as they progressed from the sixth grade to the junior year of high school
3. To evaluate the use of national stanines as a methodology in the identification of the potentially superior high school student
4. To evaluate the use of national stanines as comparable measures when different tests are used.

Achievement was measured by scores on standardized tests which were given at regular intervals throughout the school years. National modal percentile norms for all tests were converted into stanine scores, and used as comparable measures. Stanine charts were plotted for all test and sub-test scores and differentials computed.

Students achieving above or below their ability or previous achievement were identified, and the number of different students involved in the significant deviations was computed.

Forty-one ability versus achievement stanine charts were plotted, thirty-three achievement versus achievement

charts, and one ability versus ability, making a total of seventy-five stanine charts. In all, 5,461 test and sub-test scores were used, and 11,783 stanine differentials computed.

Summary charts of ability versus ability, ability versus achievement, achievement versus achievement, and a composite of all three were made.

CHAPTER IV

SUMMARY OF FINDINGS AND TENTATIVE CONCLUSIONS

1. Determination of Stanine Stability

Stanine scores permit the comparison of both individual and group performances in various phases of the educational program. Stanine stability is determined by the stanine differential obtained when two stanine scores are compared. In this study, when the differential was two or less a student was considered to show stanine stability. That is, his score on one test was within one standard deviation of his score on another test. Conversely, if a stanine differential of three or more was found between two test scores the student was considered to show a significant deviation or stanine instability. For instance, if a student scored in stanine 9 on one test, and stanine 6 on another, it was said that he showed instability between the two tests.

Stanine differentials of zero to three or more were computed between all tests discussed in this chapter. However, since this study dealt primarily with significant deviations, differentials of three or more, this chapter will be confined to an analysis of these significant deviations. Tables showing more detailed data may be found in the appendix.

2. Stability Between Mental Ability Test Scores

Stanine scores of all pupils were calculated from the national modal percentile ranks obtained in the two mental ability tests. These tests were the California Short-Form Test of Mental Maturity in the sixth grade, and the Otis Quick-Scoring Test of Mental Ability in the ninth grade. The stanine scores of 194 students were compared at both grade levels.

Table 4. Stanine Variations in the Scores of 194 Students Between the California Short-Form Test of Mental Maturity - Grade Six, and the Otis Quick-Scoring Test of Mental Ability - Grade Nine

Number and Per Cent of Students Showing Variations (1)	Stanine Variations Between Scores				
	0	1	2	3 or more	Total 0,1,2
	(2)	(3)	(4)	(5)	(6)
Number.....	62	93	29	10	184
Per Cent.....	32.0	47.9	14.9	5.2	94.8

In column (6) of Table 4 it may be seen that 184 students, or approximately 95 per cent, showed stanine differentials of two or less between mental ability scores. Of these, 62, or 32 per cent, showed no stanine variations; and 93, or 48 per cent, operated within a one-stanine range. Thus, approximately 80 per cent of the students functioned within a one-stanine limit. Only ten students showed a significant deviation.

These results suggest that a high degree of stanine stability in mental ability exists between the national modal

stanine scores of the California Short-Form Test of Mental Maturity at grade six, and the Otis Quick-Scoring Test of Mental Ability at grade nine.

3. Stability Between CTMM Scores and Achievement

The stanine scores on the California Short-Form Test of Mental Maturity, which was given in the sixth grade, were used as the ability criteria for the comparison of achievement test scores in grades six, seven, and nine.

Stability at grade six.-- The two standardized achievement test scores used by the Providence schools in the placement of all seventh-grade students for both regular and special progress classes are the sub-tests of reading comprehension and arithmetic problems of the Metropolitan Achievement Tests. For this reason, in grade six, the beginning point of this study, only the scores on these two sub-tests were recorded for comparison with ability. All three scores were available for 183 of the 200 students used in this study.

The comparison of these scores, as shown in Table 5, column (6), indicated that approximately 84 per cent of the students showed stanine stability between mental ability and achievement in both reading and arithmetic problems. These results were very similar to those reported at this level by

Corcoran ^{1/} in his study of stanine stability based on the use of local norms.

Table 5. Stanine Stability in the Scores of 183 Students Between The California Short-Form Test of Mental Maturity and the Sub-tests of Reading and Arithmetic Problems of The Metropolitan Achievement Tests - Grade Six

Name of Test	Number of Students Tested	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Reading					
Number..	183	0	30	30	153
Per Cent		0	16.4	16.4	83.6
Arithmetic Problems					
Number..	183	1	27	28	155
Per Cent		0.5	14.8	15.3	84.7
Total Test Scores					
Number..	366	1	57	58	308
Per Cent	100	0.3	15.6	15.8	84.2

Note in column (5) of Table 5 how little difference there is between the number of significant deviations in each subject. Columns (3) and (4) show that all but one of the significant deviations were minus, which indicated that about 16 per cent of the achievement test scores were below the ability

^{1/}John E. Corcoran, A Longitudinal Study to Determine the Stanine Stability of a Group's Test-Score Performance in the Elementary School, Unpublished Doctoral Dissertation, School of Education, Boston University, 1958, p. 82.

range of the students.

In general, the stanine scores of ability and achievement show a high degree of stability at the sixth-grade level.

Stability at grade seven.-- The achievement test used at this level was the Stanford Achievement Test. All six individual sub-test scores were compared with the CTMM scores.

It may be observed in column (6) of Table 6 that the stanine stability in the seventh grade ranged from 69 to 80 per cent, with a composite stability of approximately 74 per cent. Note that the highest degree of stability was found in word meaning, and the lowest in paragraph meaning and spelling. The critical ratio of 2.54 between word meaning and paragraph meaning is significant at the 5 per cent level and falls just short at the 1 per cent level. This is an interesting finding as both reading and vocabulary are usually considered to be closely related to mental ability.

Some possible explanations of the lower degree of stability found in all subjects at the seventh-grade level might be the series of adjustments which the student must make when he enters junior high school. He must adjust to a much larger school, to departmentalized work, and to many different teachers, all of which may interfere temporarily with his functioning at his maximum level of achievement. In general, there is a substantial degree of stability

Table 6. Stanine Stability in the Scores of 194 Students Between the California Short-Form Test of Mental Maturity and each Sub-test of the Stanford Achievement Test in Grade Seven

Name of Sub-Test	Number of Students Tested	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Paragraph Meaning					
Number.....	194	0	60	60	134
Per Cent....		0	31	31	69
Arithmetic Reasoning					
Number.....	191	1	43	44	147
Per Cent....		0.5	22.5	23	77
Word Meaning					
Number.....	192	3	35	38	154
Per Cent....		1.6	18.2	19.8	80.2
Arithmetic Computation					
Number.....	192	0	45	45	147
Per Cent....		0	23.4	23.4	76.6
Language Usage					
Number.....	182	1	50	51	131
Per Cent....		0.5	27.5	28	72
Spelling					
Number.....	194	0	58	58	136
Per Cent....		0	29.9	29.9	70.1
Total Test Scores					
Number....	1145	5	291	296	849
Per Cent..	100	0.4	25.4	25.8	74.2

between mental ability and achievement at the seventh-grade level, although it is considerably less than that found in the sixth grade.

Stability at grade nine.-- The same ability scores and those of an equivalent form of the achievement test used at grade seven were compared at the ninth-grade level. As Table 7 indicates, the degree of stability was considerably higher in all subjects. The range of stability varied from 79 to 95 per cent, with a composite stability of approximately 85 per cent. Column (6) shows that the pattern of stability is similar to that of the seventh grade, with the highest percentage of stability in word meaning, and the lowest in spelling. Note, however, that at this level, paragraph meaning ranks second in stability. The difference between the percentages of paragraph meaning and word meaning in the ninth grade are not significant.

The data indicate that, in general, there is a high degree of stability between ability and achievement at the ninth-grade level.

4. Significant Deviations Between CTMM Scores and Achievement

Whenever a stanine differential of three or more exists between a student's ability score and his achievement score the difference between the two scores is considered to be statistically significant. In this study this significant difference is referred to as a significant deviation. A plus deviation indicates that the student is achieving above

Table 7. Stanine Stability in the Scores of 194 Students Between the California Short-Form Test of Mental Maturity and Each Sub-test of the Stanford Achievement Test in Grade Nine

Name of Sub-test	Number of Students Tested	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Paragraph Meaning					
Number.....	193	1	17	18	175
Per Cent....		0.5	8.8	9.3	90.7
Arithmetic Reasoning					
Number.....	192	0	36	36	156
Per Cent....		0	18.8	18.8	81.2
Word Meaning					
Number.....	193	5	4	9	184
Per Cent....		2.6	2.1	4.7	95.3
Arithmetic Computation					
Number.....	192	0	36	36	156
Per Cent....		0	18.8	18.8	81.2
Language Usage					
Number.....	192	5	30	35	157
Per Cent....		2.6	15.6	18.2	81.8
Spelling					
Number.....	192	4	37	41	151
Per Cent....		2.1	19.3	21.4	78.6
Total Test Scores					
Number....	1154	15	160	175	979
Per Cent..	100	1.3	13.9	15.2	84.8

his range of ability, and a minus deviation indicates that he is achieving below the range of his ability.

In order to show a comparison of the significant deviations

between mental ability and individual subject scores in grades six, seven, and nine, Table 8 was prepared. This table also indicates, column (6), the number of different students who were involved in the significant deviations in each subject in all three grades.

Deviations in reading.-- As indicated in Tables 5, 6, and 7, only one student showed a plus deviation in reading in all three grades. Column (2) in Table 8 shows that in grade six about 16 per cent of the students were reading below their ability. This percentage almost doubled, column (3), in grade seven. By the ninth grade, column (4), it had dropped to about 9 per cent. The total number of significant deviations in reading for all three grades was 108, or approximately 19 per cent of the total test scores in reading. Note in column (6), however, that only 75 students were responsible for the 108 deviations, which indicates that 119 students worked within the range of their ability in all three grades.

The critical ratios between the grade percentages of significant deviations between ability and reading indicate that:

1. It is almost certain that a greater percentage of students will deviate significantly in reading in the seventh grade than in either the sixth (CR 3.4), or the ninth (CR 5.6).

Table 8. A Comparison of the Significant Deviations in the Achievement Scores of the Same Group of 194 Students as They Progressed from the Sixth Grade to the Ninth, When the CTMM Is Used as a Point of Reference

Name of Sub-test	Number and Per Cent of Significant Deviations			Number and Per Cent of Total Deviations	Number and Per Cent of Different Students Involved by Subject
	Grade Six	Grade Seven	Grade Nine		
(1)	(2)	(3)	(4)	(5)	(6)
Reading					
Number of Scores.....	183	194	193	570	
Deviations.	30	60	18	108	75
Per Cent...	16.4	31.0	9.3	18.9	38.7
Arithmetic Reasoning					
Number of Scores.....	183	191	192	566	
Deviations.	28	44	36	108	73
Per Cent...	15.3	23.0	18.8	19.1	37.7
Word Meaning					
Number of Scores.....	-	192	193	385	
Deviations.	-	38	9	47	45
Per Cent...	-	19.6	4.7	12.2	23.2
Arithmetic Computation					
Number of Scores.....	-	192	192	384	
Deviations.	-	45	36	81	65
Per Cent...	-	23.4	18.8	21.1	33.5
Language Usage					
Number of Scores.....	-	182	192	375	
Deviations.	-	51	35	86	69
Per Cent...	-	28	18.2	23.0	35.6
Spelling					
Number of Scores.....	-	194	192	386	
Deviations.	-	58	41	99	73
Per Cent...	-	29.9	21.4	25.6	37.6

(concluded on next page)

Table 8. (concluded)

Name of Sub-test	Number and Per Cent of Significant Deviations			Number and Per Cent of Total Deviations	Number and Per Cent of Different Students Involved by Subject
	Grade Six	Grade Seven	Grade Nine		
(1)	(2)	(3)	(4)	(5)	(6)
Total Deviations in Reading and Arithmetic Reasoning					
Number of Scores.....	366	385	385	1136	
Deviations.....	58	104	54	216	99
Per Cent...	15.8	27.0	14.0	19.0	51.0
Total Deviations in all Subjects					
Number of Scores.....	366	1145	1154	2665	
Deviations.....	58	296	175	529	148
Per Cent...	15.8	25.9	15.2	19.8	76.3

2. Although a slightly higher percentage of students will deviate in the sixth grade than in the ninth, the difference between the percentages is not significant (CR 2.08).

From this data it would appear that the reading achievement in grade seven is not the best criterion of a student's potential reading ability, although it is almost certain that the stanine stability at this level will be

between 60 and 78 per cent.

Deviations in arithmetic reasoning.-- As can be seen in Table 8, a comparison of the significant deviations in arithmetic reasoning shows that deviations in the seventh grade were approximately 8 per cent higher than in grade six, and about 4 per cent higher than in grade nine. Grade nine shows an increase of 3.5 per cent over grade six. The total number of deviations in arithmetic reasoning was 108, or about 19 per cent. Column (6) shows that 73 students were responsible for 108 deviations, which indicates that 121 students worked within the range of their ability at all three grade levels. As shown in Tables 5 and 6, one student achieved above the range of his ability in the sixth grade, and one in the seventh grade. All other deviations were minus.

The data on arithmetic reasoning suggests that although the total number of significant deviations is the same as that in reading, the deviations are more evenly distributed from grade to grade. No significant differences were found between the percentages of deviations at the three grade levels. However, the percentage of deviations in grade seven was somewhat higher than that of the other grade levels, which indicates once again that the seventh-grade achievement is not as reliable as that of the sixth grade as a criterion of a student's potential.

Total deviations in reading and arithmetic reasoning.--

Since reading and arithmetic reasoning are two of the criteria used by the Providence school in the classification of students, stanine differentials were computed between a composite score of these two sub-tests. Table 8 (concluded) shows the results of the findings which were:

1. About 16 per cent of the total test scores in the two subjects deviated significantly from ability scores in the sixth grade; 27 per cent, in grade seven; and 14 per cent, in grade nine.
2. Nineteen per cent of all scores in the two sub-tests in grades six, seven, and nine deviated significantly from ability.
3. About one half of the group showed at least one significant deviation in either reading or arithmetic reasoning between the sixth and ninth grades.
4. Approximately one half of the students worked within their ability range in grades six, seven, and nine in both reading and arithmetic reasoning.

The critical ratios obtained between the percentages of deviations at the three grade levels indicate that (1) it is almost certain that the seventh-grade scores will show greater deviation than those in either the sixth grade (CR 3.7), or the ninth (CR 4.4); and (2) the percentage of deviations in the total scores in the sixth grade will not vary significant-

ly from that in the ninth grade (CR 0.73).

Deviations in other achievement.-- Columns (3), (4), and (5) in Table 8 show the percentage of significant deviations in the other four sub-tests in grades seven and nine. Note that the percentage in word meaning was more than four times as great in the seventh grade as in the ninth. Deviations in arithmetic computation, language usage, and spelling were all slightly higher in the seventh grade than in the ninth, but no significant difference was found in the percentages at the two grade levels.

Total deviations in all subjects.-- The calculation of the sum total of stanine differentials of three or more between the scores on the ability test and the scores on all achievement sub-tests in grades six, seven, and nine yielded the following information:

1. The percentage of all significant deviations in the two sub-tests used in the sixth grade was about 16 per cent of the total number of test scores.
2. Approximately 26 per cent of the seventh-grade scores, which included all six sub-tests, showed significant deviations.
3. The percentage of deviations in all six subjects dropped to approximately 15 per cent in the ninth grade, which was within 1 per cent of the percentage of total deviations in the two sub-tests in grade

six.

4. Approximately 20 per cent of the total number of all scores recorded in grades six, seven, and nine showed significant deviations between ability and achievement.
5. The percentage of significant deviations in reading and arithmetic reasoning in grades seven was approximately the same as that of the deviations in all six subjects. This same pattern was found at the ninth-grade level.

No significant difference was found between the percentage of total deviations in reading and arithmetic reasoning in the sixth grade, and that of the total deviations in all six sub-tests in the ninth grade (CR 0.31). However, the difference between the percentage of total deviations in reading and arithmetic reasoning in grade six and that of all six subjects in grade seven was highly significant (CR 4.7). This highly significant difference was also found between the percentage of total deviations in all subjects in grade seven and that in grade nine.

These findings indicate that, in general, the percentage of the total number of significant deviations in reading and arithmetic reasoning in grade six will be about the same as that of all tests in grade nine, while the percentage of total deviations in all tests in grade seven will show a significant

increase over that in both grade six and grade nine.

6. Summary of Stanine Stability Between
CTMM Scores and Achievement

The stanine scores of the California Short-Form Test of Mental Maturity, which was given to 194 students in the sixth grade, were used as the ability criteria for the calculation of ability-achievement differentials in grades six, seven, and nine. The achievement test scores used were:

1. The sub-tests of reading and arithmetic problems of the Metropolitan Achievement Tests in grade six
2. All sub-tests of the Stanford Achievement Test in grades seven and nine.

Significant deviations, those of three or more stanines, were computed, and individual students whose scores deviated were identified in each sub-test. In all, 2,665 differentials were used.

Summary of the findings.-- The results obtained from the calculation of the ability-achievement differentials in all three grades were the following:

1. Approximately 84 per cent of the scores in the two sub-tests of reading and arithmetic problems were within the ability range of the students in the sixth grade. One student showed a plus deviation in arithmetic problems, and all other significant deviations were minus.

2. The stanine stability between ability and achievement in grade seven dropped to 73 per cent in the composite scores of reading and arithmetic reasoning, and to 74 per cent in a composite of all subjects. Approximately 98 per cent of the significant deviations indicated that the students were achieving below their ability range.
3. The percentage of the number of scores within the ability range in reading and arithmetic reasoning in grade nine increased to 86 per cent, which was within two percentage points of the sixth-grade scores. The percentage of the total number of test scores in all subjects at this level was 84.8 per cent, which was within 0.6 of a percentage point of the two sixth-grade scores which were used as the original criteria. Approximately 91 per cent of the significant deviations were minus.
4. The composite scores for all three grades indicated that 81 per cent of all scores in reading and arithmetic reasoning, and 80 per cent of all scores in all subjects fell within a two-stanine range of ability. Approximately 76 per cent of the students showed at least one significant deviation, and 24 per cent of the students worked within their ability range in all subjects in all three grades. About

96 per cent of the significant deviations indicated achievement below ability.

Table 9 shows the total stability between ability and achievement at the three grade levels, and the composite stability for all three grades.

Table 9. A Comparison of the Stability in the Scores of the Same 194 Students Between the California Short-Form Test of Mental Maturity and All Achievement Tests in Grades Six, Seven, and Nine

Name of Achievement Test	Number of Total Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Stanine Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Metropolitan Achievement Tests					
Number...	366	1	57	58	308
Per Cent.		0.3	15.6	15.8	84.1
Stanford Achievement Tests					
Grade 7					
Number...	1145	5	291	296	849
Per Cent.		0.4	25.4	25.8	74.2
Grade 9					
Number...	1154	15	160	175	979
Per Cent.		1.3	13.9	15.2	84.8
Total					
Number...	2299	20	451	471	1828
Per Cent.		0.9	19.6	20.4	79.5
Total Test Scores					
Number...	2665	21	508	529	2136
Per Cent.	100	0.8	19.1	19.9	80.2

It can be seen from the data in Table 9 that a substantial degree of stanine stability exists between ability and achievement in grades six, seven, and nine. It is interesting to note in column (6) how little the percentage of total stability in the sixth grade varies from that in the ninth although more than three times as many scores were involved in the ninth grade.

Implications of the findings.-- The results of the calculation of 2,665 stanine differentials between the sixth-grade scores on the California Short-Form Test of Mental Maturity and the scores on all achievement tests recorded in grades six, seven, and nine suggested the following:

1. The percentage of scores which fall within the ability range of the students in reading and arithmetic problems in the sixth grade is a fairly reliable indication of the percentage of scores in all achievement tests in grades six, seven, and nine which will fall within the ability range. The difference between the percentage of stability in grade six and that in all three grades is not significant.
2. There is very little chance that the percentage of stanine stability in grade six will vary significantly from that in grade nine (CR 0.31).
3. The seventh-grade differentials between ability and

achievement do not appear to be the best indication of potential scores which will fall within the ability range. It is almost a certainty that the percentage of scores which will fall within the ability range in the seventh grade will be lower than that in the sixth grade (CR 4.7), ninth grade (CR 7.1), and a composite of all three (CR 4.6).

4. The stanine stability between ability and composite achievement in reading and arithmetic reasoning at a grade level is a fairly reliable indication of the percentage of the total scores which will fall within the ability range in all six subjects at the same grade level.

7. Stability Between Otis Ability Scores and Achievement

The scores on the Otis Quick-Scoring Mental Ability Test, which was given in the ninth grade, were used as the ability criteria for the comparison of achievement scores in grades nine, ten, and eleven.

Stability at grade nine.-- The achievement test used in the first half of the ninth grade was the Stanford Achievement Test. All six sub-tests were compared with the Otis scores, as shown in Table 10. Note in column (6) that the stanine stability between ability and achievement is 90

Table 10. Stanine Stability in the Scores of 200 Students Between the Otis Test of Mental Ability and Each Sub-test of the Stanford Achievement Test in Grade Nine

Name of Sub-test	Number of Students Tested	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Paragraph Meaning					
Number...	200	1	15	16	184
Per Cent.		0.5	7.5	8.0	92.0
Arithmetic Reasoning					
Number...	197	2	16	18	179
Per Cent.		1.0	8.1	9.1	90.9
Word Meaning					
Number...	199	10	5	15	184
Per Cent.		5.0	2.5	7.5	92.5
Arithmetic Computation					
Number...	198	1	31	32	166
Per Cent.		0.5	15.7	16.2	83.8
Language Usage					
Number...	198	9	9	18	180
Per Cent.		4.5	4.5	9.0	91.0
Spelling					
Number...	198	5	28	33	165
Per Cent.		2.5	14.1	16.6	83.4
Total Test Scores					
Number...	1190	28	104	132	1058
Per Cent.	100	2.4	8.7	11.1	88.9

per cent or better in all subjects except arithmetic computation and spelling. The stability in these two subjects was approximately 83 per cent.

It is interesting to note in columns (3) and (4) that two thirds of the significant deviations between ability and achievement in word meaning were plus deviations. About the same number of students achieved above their range of ability in language as in word meaning. The plus deviations in language represented one half of the total deviations in this area. The majority of the significant deviations in all other subjects were on the minus side.

The composite stability between ability and achievement was approximately 89 per cent which indicates a high degree of stability exists between the Otis and Stanford Achievement Test Scores at the ninth-grade level.

Stability at grades nine and ten.-- From the second half of the ninth grade through the tenth grade, most of the subjects were elective. For this reason, the number of students tested varied considerably from subject to subject. However, the total number of test scores per student was approximately the same. The achievement tests used at these levels were the Cooperative Tests.

The stanine stability found between ability and achievement in the various subject areas may be seen in Table 11.

The data in column (6) indicate that 90 per cent or

Table 11. Stanine Stability in the Scores of 200 Students
Between The Otis Test of Mental Ability and
Cooperative Tests in Grades Nine and Ten

Name of Test	Number of Students Tested	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Algebra Elementary					
Number...	166	5	13	18	148
Per Cent.		3.0	7.8	10.8	89.2
Latin Elementary					
Number...	100	1	7	8	92
Per Cent.		1.0	7.0	8.0	92.0
Latin Intermed.					
Number...	86	1	30	31	55
Per Cent.		1.2	34.9	36.1	63.9
French Elementary					
Number...	81	0	16	16	65
Per Cent.		0	19.8	19.8	80.2
Reading Comp.					
Number...	184	2	8	10	174
Per Cent.		1.1	4.3	5.4	94.6
Mech. of Expression					
Number...	192	0	31	31	161
Per Cent.		0	16.1	16.1	83.9
Effect. of Expression					
Number...	178	1	23	24	154
Per Cent.		0.6	12.9	13.5	86.5
Algebra Intermed.					
Number...	60	0	6	6	54
Per Cent.		0	10.0	10.0	90.0
Biology Elementary					
Number...	83	8	2	10	73
Per Cent.		9.6	2.4	12.0	88

(concluded on next page)

Table 11. (concluded)

Name of Test	Number of Students Tested	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
European History					
Number...	22	0	2	2	20
Per Cent.		0	9.9	9.9	90.0
Total Test Scores					
Number...	1152	18	138	156	996
Per Cent.	100	1.6	11.9	13.5	86.5

better of the students tested showed stanine stability between ability and achievement in the areas of elementary Latin, reading, intermediate algebra, and European history. A stability of 80 per cent or above was found in all areas except intermediate Latin. This low stability between ability and intermediate Latin is interesting as it represents a significant drop from that at the elementary level. This difference between the two levels warrants further study as it may be an indication that the content of this test is not well adapted to the Providence course of study at this level.

In general, a substantial to high degree of stability exists between ability and the Cooperative Tests in all areas except intermediate Latin.

Stability at grade eleven.-- The Iowa Tests of Educational Development were the tests used as achievement measures in the eleventh grade. However, unlike the other achievement tests which were administered to all students, the ITED was administered to only 121 of the 200 students. A limited budget, plus the fact that the tests had to be administered during a two-day school holiday, made it impossible for all students to participate in the ITED. All students who were tested at this level had an ability stanine of six or higher.

The results of the calculation of the stanine differentials between ability and the ITED scores are tabulated in Table 12. Column (6) shows that the stability was approximately 92 per cent or above in all but three of the tests. The lowest degree of stability was found in the test on natural science background. The significantly high percentage of students working below their ability range in this area, as shown in column (4), suggests that the course of study in Providence, prior to the junior year, may be weak in science. This weakness was especially prevalent among the students who were enrolled in a classical school, as 50 per cent of those tested scored below their ability range.

It is interesting to note, column (6), that despite the low scores in natural science background, only one student failed to score within his ability range on the composite

Table 12. Stanine Variations in the Scores of 121 Selected Students Between The Otis Test of Mental Ability and The Iowa Tests of Educational Development in the Eleventh Grade

Subject	Number of Students Tested	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Stanine Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Social Concepts					
Number....	121	1	19	20	101
Per Cent..		0.8	15.7	16.5	83.5
Natural Science					
Number....	121	0	43	43	78
Per Cent..		0	35.5	35.5	64.5
Correctness of Express.					
Number....	120	2	1	3	117
Per Cent..		1.7	0.83	2.5	97.5
Quantitative Thinking					
Number....	120	1	8	9	111
Per Cent..		0.83	6.7	7.6	92.4
Reading Soc. Studies					
Number....	120	0	8	8	112
Per Cent..		0	6.7	6.7	93.3
Reading Nat. Sciences					
Number....	120	0	10	10	110
Per Cent..		0	8.3	8.3	91.7
Reading Lit. Materials					
Number....	121	0	7	7	114
Per Cent..		0	5.8	5.8	94.2
Vocabulary					
Number....	121	2	3	5	116
Per Cent..		1.7	2.5	4.1	95.9
Source Information					
Number....	121	0	13	13	108
Per Cent..		0	10.7	10.7	89.3

(concluded on next page)

Table 12. (concluded)

Subject	Number of Students Tested	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Stanine Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Composite Score					
Number....	120	0	1	1	119
Per Cent..		0	0.8	0.8	99.2
Total Test Scores					
Number....	1205	6	113	119	1086
Per Cent..	100	0.5	9.4	9.9	90.1

score, and 90 per cent of all test scores showed stanine stability between ability and achievement.

Total stability at grades nine, ten, and eleven.-- In order to show a comparison of the stanine stability between ability and achievement in grades nine, ten, and eleven, Table 13 was prepared. As may be observed in column (6), the total scores in each achievement test battery showed a stability of approximately 87 per cent or higher. The composite stability was approximately 89 per cent, which indicates a high degree of stability exists between ability and achievement in grades nine to eleven.

The difference between the percentage of scores within the ability range on all achievement tests in grades nine

Table 13. A Comparison of the Stanine Stability in the Scores of the Same 200 Students Between the Otis Test of Mental Ability and All Achievement Tests in Grades Nine to Eleven

Name of Achievement Test	Number of Total Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Stanford Achievement Tests					
Number.....	1190	28	104	132	1058
Per Cent.....		2.4	8.7	11.1	88.9
Cooperative Tests					
Number.....	1152	18	138	156	996
Per Cent.....		1.6	11.9	13.5	86.5
Iowa Tests of Educational Development					
Number.....	1205	6	113	119	1086
Per Cent.....		0.5	9.4	9.9	90.1
Total Test Scores					
Number.....	3547	52	355	407	3140
Per Cent.....	100	1.5	10.0	11.5	88.5

to eleven and that of the sixth-grade criteria scores was not significant at the 1 per cent level.

8. Different Students Involved in Significant

Deviations Between Ability Scores and Achievement

The studies concerning capacity - achievement measures found in the research showed the number and per cent of significant deviations between ability and achievement scores

in various grades and subjects, but did not indicate whether these deviations were scored by the same students from grade to grade, or whether different students were responsible for the deviations in different grades. This study has attempted to determine how many different students showed significant deviations, and at what grade level the first deviation occurred. This data is tabulated in Table 14.

Deviations in reading.--- Column (2) in Table 14 shows that in grade six 30 students deviated significantly between ability and reading achievement. In grade seven, 40 additional students showed significant deviations. Since the total number of significant deviations in reading in grade seven was 60, as shown in column (3), of which 40 were first deviations, this indicated that 18 students who showed a deviation in grade six were still reading below their ability range in grade seven. In grade nine there were 16 deviations, 9 of which involved first deviations, which indicated that 7 students who had been below ability in reading in grade six or seven were still achieving below ability in grade nine. Of these, five had deviated in grades six and seven. Three students deviated for the first time in grade ten.

The identification of different students involved in significant stanine deviations between ability and reading achievement indicated the following:

1. About 15 per cent of the students showed deviations

Table 14. Number of Different Students, In Group of 200 Students, Involved in Significant Deviations Between Ability and Achievement in Grades Six, Seven, Nine, and Ten

Number and Per Cent of Different Students Involved in Total Deviations	Grade Six	Grade Seven	Grade Nine	Grade Ten	Total Number of Different Students Involved
(1)	(2)	(3)	(4)	(5)	(6)
Reading					
Total deviations...	30	60	16	10	116
Number of students.	30	42	9	3	84
Per Cent.....	15.0	21.0	4.5	1.5	42.0
Arithmetic Reasoning and Algebra					
Total deviations...	28	44	36	6	114
Number of students.	28	33	12	2	75
Per Cent.....	14.0	16.5	6.0	1.0	37.5
Word Meaning					
Total deviations...	-	38	15	-	53
Number of students.	-	38	11	-	49
Per Cent.....	-	19.0	5.5	-	24.5
Arithmetic Computation					
Total deviations...	-	45	32	-	77
Number of students.	-	45	22	-	67
Per Cent.....	-	22.5	11.0	-	33.5
English					
Total deviations...	-	51	18	55	124
Number of students.	-	51	13	42	106
Per Cent.....	-	25.5	6.5	21.0	53.0
Spelling					
Total deviations...	-	58	33	-	91
Number of students.	-	58	15	-	73
Per Cent.....	-	29.0	7.5	-	36.5
Languages					
Total deviations...	-	-	24	31	55
Number of students.	-	-	24	19	43
Per Cent.....	-	-	12.0	9.5	21.5
Biology and History					
Total deviations...	-	-	-	12	12
Number of students.	-	-	-	12	12
Per Cent.....	-	-	-	6.0	6.0
All subjects					
Total deviations...	58	296	174	114	642
Number of students.	53	83	23	14	173
Per Cent.....	26.5	41.5	11.5	7.0	86.5

in grade six.

2. By grade seven, 21 per cent more had become involved, which made a total of approximately 36 per cent.
3. Six per cent of the students showed a first deviation in reading after the seventh grade.
4. Approximately 42 per cent of the group showed a significant deviation in reading, while 58 per cent of the group worked within a two-stanine range in all four grades.
5. Approximately 86 per cent of all students who deviated in reading did so by the seventh grade.

Deviations in arithmetic reasoning.-- In Table 14 it may be seen that 28 students showed significant deviations between ability and arithmetic reasoning in grade six. In grade seven, 33 additional students deviated; in grade nine, 12, and in grade ten, 2. The total number of deviations in grade seven was 44, which indicated that 11 of the students deviated in both the sixth grade and the seventh. The total of 36 deviations in grade nine indicated that 24 of the students had already deviated in a previous grade. Two students deviated for the first time in grade ten.

The above data yielded the following information:

1. About 14 per cent of the group were working below ability in arithmetic reasoning in grade six.
2. Thirty-one per cent of the group had shown at least

one deviation by the seventh grade.

3. Six per cent of the group deviated for the first time in the ninth grade, and one per cent in the tenth.
4. Approximately 38 per cent of the students showed at least one deviation in the four grades; and approximately 62 per cent showed no significant deviations.
5. About 81 per cent of all students who deviated in arithmetic reasoning did so by the seventh grade.

Deviations in reading and arithmetic reasoning.-- When the test scores on both the reading and arithmetic reasoning were combined, the total number of different students who showed significant deviations was 53, or about 27 per cent in the sixth grade. Forty-six additional students showed deviations in the seventh grade, and fifteen in the ninth grade. Four deviated for the first time in these two subjects in grade ten. The total percentage of students showing deviations in these two subjects was approximately 60 per cent for all four grades. Approximately eighty-three per cent of the deviations occurred by the seventh grade.

Deviations in English.-- Column (6) of Table 14 shows that the largest number of different students who deviated significantly in any one subject was in English. Note in column (5) that this was the only subject which showed a

large number of significant deviations at the tenth-grade level. This is an interesting observation and might warrant further investigation which was not within the scope of this study.

Deviations in other subjects.-- Approximately 25 per cent of the students deviated in word meaning. Of these, 19 per cent deviated in the seventh grade.

Approximately 34 per cent of the group deviated in arithmetic computation, 23 per cent in the seventh grade.

In spelling, about 37 per cent of the students showed significant deviations. Twenty-nine per cent of them did so in the seventh grade.

Approximately 29 per cent of the students deviated in elective subjects.

Deviations in all subjects.-- About 87 per cent of all the students showed at least one significant deviation, and 13 per cent of the group worked within the range of their ability in all subjects in grades six, seven, nine, and ten. The greatest number of first deviations, 49 per cent, occurred in the seventh grade. Twenty students deviated only once, fifteen of them at the seventh-grade level. Four students showed all plus deviations.

The preceding data suggest that the greatest degree of stanine instability between ability and achievement for the individual student is found at the seventh-grade level.

9. Summary of Stanine Stability Between Ability and Achievement

The California Test of Mental Maturity and The Otis Test of Mental Ability were the two ability tests used in this study. The former test was administered in the sixth grade and the latter in the ninth grade. The scores on the CTMM were projected through the ninth grade, and the Otis scores from the ninth to the eleventh. In all, 6,212 stanine differentials were computed between ability and achievement test scores from grades six to eleven. The percentage of the number of sixth grade stanine scores in the sub-tests of reading and arithmetic problems which fell within the ability range was used as a criterion to determine if there would be any significant deviation between this percentage and those of all other achievement test scores within ability range as the students progressed from the sixth grade to the eleventh.

Results of the findings.--- The results of the calculations of ability-achievement differentials which are shown in Table 15 were as follows:

1. In the sixth grade, approximately 84 per cent of the achievement scores in both reading and arithmetic problems were within a two-stanine range of the students' ability.
2. In grade seven, the percentage of the combined scores within ability range in reading and arithmetic

reasoning dropped to 73, and the percentage of the composite score of all achievement tests was about 74 per cent.

3. In grade nine, when the CTMM was used as a point of reference, the percentage of the composite score of all achievement test scores within the ability range was about 85. When the Otis test was used as a point of reference the percentage was a few points higher, 89. The difference between the standard errors of the two percentages was 2.7, which indicates that the difference is probably a significant one.
4. The percentage of the total number of CTMM differentials between achievement scores which fell within a two-stanine range was about 80 per cent.
5. The percentage of the sum total of differentials between the Otis and all achievement test scores in grades nine to eleven which were two or less was about 89 per cent. In the Cooperative Tests, which were given in grades nine and ten, approximately 87 per cent of all scores were within the ability range. In the ITED, which was given to 121 eleventh grade students who scored in the sixth stanine or above on the CTMM, 90 per cent of all scores were within a two-stanine range.

Table 15. A Comparison of the Total Stanine Variations in all Ability-Achievement Scores of the Same 200 Students As They Progressed from the Sixth Grade to the Eleventh

Names of Ability and Achievement Tests	Number of Total Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Stanine Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
California Test of Mental Maturity Metropolitan Achievement Tests					
Number....	366	1	57	58	308
Per Cent..		0.3	15.6	15.8	84.1
Stanford Achievement Tests					
Grade 7					
Number....	1145	5	291	296	849
Per Cent..		0.4	25.4	25.8	74.2
Grade 9					
Number....	1154	15	160	175	979
Per Cent..		1.3	13.9	15.2	84.8
Total					
Number....	2299	20	451	471	1828
Per Cent..		0.9	19.6	20.4	79.5
Total CTMM					
Number....	2665	21	508	529	2136
Per Cent..		0.8	19.1	19.9	80.2
Otis Test of Mental Ability Stanford Achievement Tests-Gr. 9					
Number....	1190	28	104	132	1058
Per Cent..		2.4	8.7	11.1	88.9
Cooperative Tests					
Number....	1152	18	138	156	996
Per Cent..		1.6	11.9	13.5	86.5

(concluded on next page)

Table 15. (concluded)

Names of Ability and Achievement Tests	Number of Total Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Stanine Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
ITED					
Number....	1205	6	113	119	1086
Per Cent..		0.5	9.4	9.9	90.1
Total-Otis					
Number....	3547	52	355	407	3140
Per Cent..		1.5	10.0	11.5	88.5
Total Test Scores					
Number....	6212	73	863	936	5276
Per Cent..	100	1.2	13.9	15.1	84.9

6. The percentage of the composite number of all ability-achievement differentials which fell within a two-stanine range over a five-year period was about 85.

Different students involved in significant deviations.--

In order to determine whether the same students consistently deviated significantly in one or more subjects each student showing a stanine deviation of three or more was identified throughout the study. The results of this identification were as follows:

1. In reading, about 42 per cent of the students were responsible for one or more significant deviations, while about 58 per cent of the students showed

stanine stability throughout the entire period.

2. In all subjects, the largest number of students, approximately 42 per cent, showed their first, and for many their only, significant deviation in the seventh grade.
3. About 19 per cent of the students deviated significantly for the first time after the seventh grade.
4. In all, about 87 per cent of the group showed at least one significant deviation throughout the five year period, while 13 per cent of the students worked within a two-stanine range in all subjects over the same period of time.

Implications of findings.-- The findings on the analysis of the data suggested the following:

1. The percentage of the number of sixth-grade ability-achievement differentials within a two-stanine range in both reading and arithmetic problems is a fairly reliable index of the percentage of the number of differentials on all achievement tests over a five-year period which will fall within a two-stanine range of ability. The critical ratio between the two percentages indicated that there is little chance that the difference will be significant.
2. The percentage of ability-achievement differentials which fell within a two-stanine range showed a ten-

point drop in the seventh grade. Calculation of the critical ratios between the percentages of total stanine stability at all grade levels indicated that it is almost certain that the seventh-grade scores will show a lower degree of stability than those at any other level. Therefore, the data indicate that the percentage of seventh-grade achievement scores within the ability range is not the best indication of the percentage of the potential number of ability-achievement scores which will fall within a two-stanine range as the student progresses through school.

3. The percentage of the number of ninth-grade ability-achievement differentials within a two-stanine range may be slightly higher when the Otis Test is used as the ability test than when the CTMM is used. This may be due to the fact that the Otis Test was given in the ninth grade and the CTMM in the sixth. An additional factor might be that the Otis showed a high correlation with The Stanford Achievement Test at the time of standardization.
4. It is almost certain that 70 per cent of the students who achieve below ability on any test will do so by the seventh grade.
5. More students will show under-achievement in the

seventh grade than at any other level.

10. Stability In Achievement Test Scores

At Different Grade Levels

Stanine stability in achievement over a period of five years was determined through the computation of stanine differentials of test scores in the same subjects as the students progressed from the sixth grade to the eleventh. Various tests and grade levels were used as points of reference to study the stability from one grade to another as well as the five-year stability.

Stability when grade-six achievement is used as a point of reference.-- The sixth-grade scores on the sub-tests of reading and arithmetic problems of the Metropolitan Achievement Tests were the stanine scores used as the first points of reference. Stanine differentials were calculated between these scores and all other achievement scores recorded in these two subjects from the seventh to the eleventh grades.

Table 16 shows the stanine stability in reading between the sixth-grade achievement and each subsequent level tested. As can be seen in column (6), the lowest stability, 85.9 per cent, was found in the seventh grade. All other levels indicated a stability of 90 per cent or higher. Note in columns (3) and (4) that the seventh grade was the only level at which the percentage of minus deviations exceeded that of the plus deviations. In general, a high degree of stability

exists between the stanine scores in reading in the sixth grade and those at other grade levels.

Table 16. A Comparison of the Stanine Stability in the Reading Performance of the Same Group of 184 Students in Grades Seven, Nine, Ten, and Eleven When Grade-Six Achievement Is Used as a Point of Reference

Name of Test and Grade Tested	Number of Students Tested	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Stanford Achievement Test					
Grade 7					
Number....	184	5	21	26	158
Per Cent..		2.7	11.4	14.1	85.9
Grade 9					
Number....	184	12	6	18	166
Per Cent..		6.5	3.3	9.8	90.2
Cooperative Test					
Grade 10					
Number....	169	5	3	8	161
Per Cent..		3.0	1.8	4.7	95.3
ITED-Test 7					
Grade 11					
Number....	112	5	3	8	104
Per Cent..		4.5	2.7	7.1	92.9
Total Test Scores					
Number....	649	27	33	60	589
Per Cent..	100	4.2	5.1	9.2	90.8

The stanine stability between the sixth-grade scores in

arithmetic problems and all subsequent achievement scores which involved mathematical reasoning is shown in Table 17.

Table 17. A Comparison of the Stanine Stability of Test Scores In Arithmetic Reasoning of the Same Group of 182 Students When Grade-Six Achievement Is Used as a Point of Reference

Name of Test	Grade and Number of Students Tested	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Stanford Achievement	Grade Seven				
Number.....	181	7	15	22	159
Per Cent...		3.9	8.3	12.2	87.8
Stanford Achievement	Grade Nine B				
Number.....	182	9	12	21	161
Per Cent...		4.9	6.6	11.5	88.5
Cooperative Algebra	Grade Nine A				
Number.....	153	17	4	21	132
Per Cent...		11.1	2.6	13.7	86.3
ITED-Test 4					
Number.....	111	8	3	11	100
Per Cent...		7.2	2.7	9.9	90.1
Total Test Scores					
Number.....	627	41	34	75	552
Per Cent...	100	6.5	5.4	12.0	88.0

As indicated in column (6), the stability ranged from 86 to 90 per cent with a composite stability of 88 per cent. The lowest degree of stability was found in algebra, but note

in columns (3) and (4) that approximately four fifths of the deviations indicated plus achievement.

Table 18 shows the stability between the composite scores in reading and arithmetic problems in the sixth grade and those at other grade levels.

Table 18. A Comparison of the Stanine Stability in the Composite Scores of Reading and Arithmetic Reasoning of the Same Group of 184 Students When Grade-Six Achievement Is Used as a Point of Reference

Name of Test	Grade and Number of Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Stanford Achievement	Grade Seven				
Number.....	365	12	36	48	317
Per Cent...		3.3	9.9	13.2	86.8
Stanford Achievement	Grade Nine				
Number.....	366	21	18	39	327
Per Cent...		5.7	4.9	10.7	89.3
Cooperative Achievement	Grades Nine and Ten				
Number.....	322	22	7	29	293
Per Cent...		6.8	2.2	9.0	91.0
ITED Tests 4 and 7	Grade Eleven				
Number.....	223	13	6	19	204
Per Cent...		5.8	2.7	8.5	91.5
Total Test Scores					
Number.....	1276	68	67	135	1141
Per Cent...	100	5.3	5.3	10.6	89.4

As can be seen in column (6), when sixth-grade achievement was used as a point of reference the stanine stability of the composite scores in reading and arithmetic reasoning ranged from 86.8 to 91.5 per cent. The lowest degree of stability was at the seventh-grade level, although no significant differences were found between the percentages at any grade level.

These results indicate that a high degree of stability exists between sixth-grade composite scores in reading and arithmetic problems, and all subsequent tests in these areas as students progress through high school.

Stability when grade-seven achievement is used as a point of reference.-- The stanine scores on the five sub-tests of the Stanford Achievement Test were the points of reference used at this level.

Table 19 shows the results of the calculations of the differentials in reading. As indicated in column (6), the stability ranged from 77 to 79.6 per cent with a composite stability of 81 per cent. This was significantly lower than that found when the sixth-grade achievement was used as a point of reference. Note, however, in columns (3) and (4) that approximately 95 per cent of the deviations indicated plus achievement. These findings suggest that reading achievement tends to be lower in the seventh grade than at other grade levels.

Table 19. A Comparison of the Stanine Stability in the Reading Performance of the Same Group of 196 Students in Grades Nine to Eleven When Seventh-Grade Achievement Is Used as a Point of Reference

Name of Test	Grade and Number of Students Tested	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Stanford Achievement Test	Grade Nine				
Number.....	196	38	2	40	156
Per Cent...		19.4	1.0	20.4	79.6
Cooperative Test	Grade Ten				
Number.....	181	24	2	26	155
Per Cent...		13.3	1.1	14.4	85.6
ITED Test 7	Grade Eleven				
Number.....	119	26	1	27	92
Per Cent...		21.8	0.8	22.7	77.3
Total Test Scores					
Number.....	496	88	5	93	403
Per Cent...	100	17.9	1.0	18.8	81.2

The stability between seventh-grade achievement in arithmetic reasoning and that of achievement in this area at other grade levels ranged from 79 to 90 per cent with a composite stability of 85 per cent. As shown in columns (3) and (4) of Table 20, the majority of the deviations indicated plus achievement. Although considerable stability was found between seventh-grade achievement in arithmetic

reasoning and that at other levels the data suggest that students tend to achieve at a lower level in the seventh grade.

Table 20. A Comparison of the Stanine Stability in the Test Scores in Arithmetic Reasoning of the Same Group of 193 Students When the Seventh-Grade Achievement Is Used as a Point of Reference

Name of Test	Grade and Number of Students Tested	Number, Per Cent, and Direction of Stanine Variations in Arithmetic Reasoning			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Stanford Achievement Test	Grade Nine				
Number.....	180	13	5	18	162
Per Cent....		7.2	2.8	10.0	90.0
Cooperative Elementary Algebra	Grade Nine				
Number.....	162	29	5	34	128
Per Cent....		17.9	3.1	21.0	79.0
ITED Test 4	Grade Eleven				
Number.....	117	15	2	17	100
Per Cent....		12.8	1.7	14.5	85.5
Total Test Scores					
Number.....	459	57	12	69	390
Per Cent....	100	12.4	2.6	15.0	85.0

The results of the calculation of the stanine differentials between seventh-grade English scores and those at other grade levels are shown in Table 21. The stability in English

Table 21. A Comparison of the English Performance of the Same Group of 183 Students In Grades Nine to Eleven When Seventh-Grade Achievement Is Used As a Point of Reference

Name of Test	Grade and Number of Students Tested	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Stanford Achievement	Grade Nine				
Number.....	183	26	3	29	154
Per Cent...		14.2	1.6	15.8	84.2
Cooperative Mechanics of Expression	Grade Ten B				
Number.....	177	7	24	31	146
Per Cent...		4.0	13.6	17.6	82.5
Cooperative Effectiveness of Expression	Grade Ten A				
Number.....	165	26	13	39	126
Per Cent...		15.8	7.9	23.7	76.3
ITED Test 3	Grade Eleven				
Number.....	107	20	3	23	84
Per Cent...		18.7	2.8	21.5	78.5
Total Test Scores					
Number.....	632	79	43	122	510
Per Cent...	100	12.5	6.8	19.3	80.7

ranged from 76 to 84 per cent with a composite stability of approximately 81 per cent. The lowest degrees of stability were found in the second half of the tenth grade and in the eleventh grade. As may be seen in columns (3) and (4), the same pattern of plus achievement accounted for the majority

of deviations. Note, however, that in the first half of the tenth grade, a majority of the deviations represented minus achievement. This is the first time that this pattern has been found.

The stanine stability between the vocabulary scores in grade seven and those in grades nine and eleven varied from 86.7 to 89.8 with a total stability of 87.9 per cent. All significant deviations indicated plus achievement as shown in Table 22.

Table 22. A Comparison of the Stanine Stability in the Vocabulary Scores of the Same Group of 195 Students in Grades Nine to Eleven When Seventh-Grade Achievement Is Used As a Point of Reference

Name of Test	Grade and Number of Students Tested	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Stanford Achievement	Grade Nine				
Number.....	195	26	0	26	169
Per Cent...		13.3	0	13.3	86.7
ITED Test 8					
Number.....	118	12	0	12	106
Per Cent...		10.2	0	10.2	89.8
Total Test Scores					
Number.....	313	38	0	38	275
Per Cent...	100	12.1	0	12.1	87.9

Only the ninth-grade scores were available for comparison with the seventh-grade scores in the areas of spelling and arithmetic computation. Test scores from the Stanford Achievement Test were used in both subjects. Table 23 indicates that the stability was approximately the same in both subjects. Note, however, in columns (3) and (4) that the deviations in arithmetic computation were almost evenly divided between plus and minus achievement, while those in spelling were practically all plus.

Table 23. A Comparison of the Performance of the Same Group of 195 Students in Arithmetic Computation and Spelling in the Ninth Grade When Seventh-Grade Achievement Is Used As a Point of Reference

Name of Test	Number of Students Tested	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Arithmetic Computation					
Number.....	180	9	10	19	161
Per Cent...		5.0	5.6	10.6	89.4
Spelling					
Number.....	195	19	2	21	174
Per Cent...		9.7	1.0	10.8	89.2
Total Test Scores					
Number.....	375	28	12	40	335
Per Cent...	100	7.5	3.2	10.7	89.3

In general a high degree of stability exists between

the seventh and ninth-grade scores in the areas of arithmetic computation and spelling.

When the seventh-grade achievement scores were used as points of reference for comparison with all subsequent achievement scores the stability in each subject ranged from 80.7 to 89.3 with a composite stability in all subjects of approximately 84 per cent. The highest stability was found in spelling, and the lowest in English. Columns (3) and (4) in Table 24 show that approximately four fifths of the scores which deviated significantly from seventh-grade achievement indicated plus achievement. The difference between the percentage of reading and arithmetic reasoning scores which fell within a normal stanine range, and that of all test scores in all subjects within a two-stanine range was not significant.

The use of seventh-grade data as points of reference suggests that the seventh-grade achievement record does not appear to be the best indication of potential achievement as it tends to be somewhat lower in all subjects than at other grade levels. Although a substantial degree of stability was found between achievement at the seventh-grade level and that at other grade levels, it was significantly lower than that found when the sixth-grade achievement was used as a point of reference.

Stability when grade-nine achievement is used as a point

Table 24. A Comparison of the Stanine Stability in the Performance of the Same Group of 200 Students on All Subsequent Achievement Tests When Grade-Seven Achievement Is Used As a Point of Reference

Subject	Number of Total Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Reading					
Number.....	496	88	5	93	403
Per Cent...		17.9	1.0	18.8	81.2
Arithmetic Reasoning					
Number.....	459	57	12	69	390
Per Cent...		12.4	2.6	15.0	85.0
Total					
Number.....	955	145	17	162	793
Per Cent...		15.2	1.8	17.0	83.0
English					
Number.....	632	79	43	122	510
Per Cent...		12.5	6.8	19.3	80.7
Vocabulary					
Number.....	313	38	0	38	275
Per Cent...		12.1	0	12.1	87.9
Arithmetic Computation					
Number.....	180	9	10	19	161
Per Cent...		5.0	5.6	10.6	89.4
Spelling					
Number.....	195	19	2	21	174
Per Cent...		9.7	1.0	10.8	89.3
Total Test Scores					
Number.....	2275	290	72	362	1913
Per Cent...	100	12.7	3.2	15.9	84.1

of reference.-- The ninth-grade scores on the sub-tests of the Stanford Achievement Test were used as points of reference to determine the stanine stability in achievement from the ninth to the eleventh grades.

A very high degree of stability in reading was found from the ninth grade to the eleventh as shown in Table 25. A composite total of the scores in both grades indicated that slightly less than 4 per cent of the scores deviated significantly, and were about equally divided between plus and minus achievement.

Table 25. A Comparison of the Stanine Stability in the Reading Performance of the Same Group of 200 Students in Grades Ten and Eleven When Ninth-Grade Achievement Is Used As a Point of Reference

Name of Test	Grade and Number of Students Tested	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Cooperative Reading Comprehension	Grade Ten				
Number.....	184	3	3	6	178
Per Cent...		1.6	1.6	3.3	96.7
ITED Test 7	Grade Eleven				
Number.....	121	3	2	5	116
Per Cent...		2.5	1.7	4.1	95.9
Total Test Scores					
Number.....	305	6	5	11	294
Per Cent...	100	2.0	1.6	3.6	96.4

The stanine scores on the sub-test of arithmetic reasoning of the Stanford Achievement Test in the first half of the ninth grade and those on the Cooperative algebra test in the second half of the ninth grade showed a substantial degree of stability. As indicated in Table 26, the majority of the significant deviations represented plus achievement. The stability between the scores in arithmetic reasoning in grade nine and those in quantitative thinking in grade eleven was very high with less than 7 per cent of the scores showing significant deviations.

Table 26. A Comparison of the Stanine Stability in the Test Scores in Arithmetic Reasoning of the Same Group of 193 Students in Grades Nine A and Eleven When Nine B Achievement Is Used As a Point of Reference

Name of Test	Grade and Number of Students Tested	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Cooperative Elementary Algebra	Grade Nine A				
Number.....	165	25	6	31	134
Per Cent...		15.2	3.6	18.8	81.2
ITED Test 4	Grade Eleven				
Number.....	120	6	2	8	112
Per Cent...		5.0	1.7	6.7	93.3
Total Test Scores					
Number.....	285	31	8	39	246
Per Cent...	100	10.9	2.8	13.7	86.3

The stanine differentials between the ninth-grade scores on the language usage test of the Stanford Achievement Test and those of all subsequent tests in English are shown in Table 27.

Table 27. A Comparison of the Stanine Stability in the Test Scores in English of the Same Group of 200 Students in Grades Nine to Eleven When Ninth-Grade Achievement Is Used As a Point of Reference

Name of Test	Grade and Number of Students Tested	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Cooperative Mechanics of Expression	Grade Ten B				
Number.....	191	8	37	45	146
Per Cent...		4.2	19.4	23.6	76.4
Cooperative Effectiveness of Expression	Grade Ten A				
Number.....	177	8	19	27	150
Per Cent...		4.5	10.7	15.2	84.8
ITED Test 3	Grade Eleven B				
Number.....	120	8	2	10	110
Per Cent...		6.7	1.7	8.4	91.6
Total Test Scores					
Number.....	488	24	58	82	406
Per Cent...	100	4.9	11.9	16.8	83.2

Note in column (6) that the range of stability varies from 76.4 to 91.6 per cent. This is the widest range of

variability found in achievement stability in any area. It is interesting to note that the pattern of minus achievement found between the seventh and tenth-grade scores in English exists at this level also. This finding may warrant further investigation.

Of the 121 students who took the ITED in grade eleven only two deviated significantly from their ninth-grade vocabulary scores. Both of these deviations indicated plus achievement.

Table 28. A Comparison of the Stanine Stability in the Performance of the Same Group of 200 Students on All Subsequent Achievement Tests When Grade-Nine Achievement Is Used As a Point of Reference

Subject	Number of Total Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Reading					
Number.....	305	6	5	11	294
Per Cent.....		2.0	1.6	3.6	96.4
Arithmetic Reasoning					
Number.....	285	31	8	39	246
Per Cent.....		10.9	2.8	13.7	86.3
Total					
Number.....	590	37	13	50	540
Per Cent.....		6.3	2.2	8.5	91.5
English Usage					
Number.....	488	24	58	82	406
Per Cent.....		4.9	11.9	16.8	83.2
Vocabulary					
Number.....	121	2	0	2	119
Per Cent.....		1.7	0	1.7	98.3
Total Test Scores					
Number.....	1199	63	71	134	1065
Per Cent.....	100	5.3	5.9	11.2	88.8

When the ninth-grade achievement on the five sub-tests of the Stanford Achievement Test was used as a point of reference for comparison with all subsequent achievement tests the stanine stability ranged from 83.2 to 98.3 per cent with a composite stability in all subjects of approximately 89 per cent. It may be seen in Table 28 that English showed the lowest degree of stability as well as the highest percentage of minus achievement. In general, it may be said, that a substantial to high degree of stability exists between ninth-grade achievement and all subsequent achievement.

Stability in grades ten and eleven.-- The Cooperative Tests, two in English and one in reading, were the reference points used in determining stability in achievement from grade ten to eleven. As indicated in column (6) of Table 29, the stability in English varied from 83 to 86 per cent with a composite stability of 84 per cent. It is interesting to note the comparatively high percentage of plus achievement at each level.

Of the 113 scores available in eleventh-grade reading, only three indicated significant deviations from tenth-grade scores.

The composite scores in English and reading in grades ten and eleven showed a stability of approximately 87 per cent with less than 3 per cent of the deviations indicating

Table 29. A Comparison of the Stanine Stability in the Achievement Test Scores of the Same Group of 178 Students in Grades Ten and Eleven

Subject and Grades Tested	Number of Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
English Ten B to Ten A					
Number.....	171	19	10	29	142
Per Cent.....		11.1	5.8	16.9	83.1
English Ten B to Eleven					
Number.....	114	19	0	19	95
Per Cent.....		16.7	0	16.7	83.3
Total English Scores					
Number.....	285	38	10	48	237
Per Cent.....	100	13.3	3.5	16.8	83.2
English Ten A to Eleven					
Number.....	115	14	2	16	99
Per Cent.....	100	12.2	1.7	13.9	86.1
Total English Scores					
Number.....	400	52	12	64	336
Per Cent.....	100	13.0	3.0	16.0	84.0
Reading Ten B to Eleven					
Number.....	113	1	2	3	110
Per Cent.....	100	0.9	1.8	2.7	97.3
Total Test Scores					
Number.....	513	53	14	67	446
Per Cent.....	100	10.3	2.8	13.1	86.9

minus achievement. The pattern of a high percentage of plus differentials between ten B English scores and subsequent

IV 58

achievement was the reverse of that found between ten B and previous achievement, which suggests that students tend to score lower in English at the tenth-grade level than at either the ninth or eleventh-grade level.

11. Significant Deviations in Achievement and Number of Different Students Involved at Different Grade Levels

In order to show a comparison of the significant deviations in achievement when different grade levels were used as points of reference, tables were prepared for each subject. These tables also indicate the number of different students who were responsible for the deviations.

Deviations in reading.-- Table 30 shows the significant deviations in reading scores between the different grades.

Table 30. A Comparison of the Stanine Variations of Three or More in the Reading Performance of the Same Group of 200 Students As They Progressed From the Sixth Grade to the Eleventh

Grades Tested	Number of Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Different Students Involved
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Six to Seven					
Number.....	184	5	21	26	26
Per Cent.....		2.7	11.4	14.1	13.0
Six to Nine					
Number.....	184	12	6	18	12
Per Cent.....		6.5	3.3	9.8	6.0

(continued on next page)

Table 30. (continued)

Grades Tested	Number of Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Different Students Involved
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Six to Ten					
Number.....	169	5	3	8	4
Per Cent.....		3.0	1.8	4.7	2.0
Six to Eleven					
Number.....	112	5	3	8	3
Per Cent.....		4.5	2.7	7.1	1.5
Total					
Number.....	649	27	33	60	45
Per Cent.....	100	4.2	5.1	9.2	22.5
Seven to Nine					
Number.....	196	38	2	40	40
Per Cent.....		19.4	1.0	20.4	20.0
Seven to Ten					
Number.....	181	24	2	26	8
Per Cent.....		13.3	1.1	14.4	4.0
Seven to Eleven					
Number.....	119	26	1	27	7
Per Cent.....		21.8	0.8	22.7	3.5
Total					
Number.....	496	88	5	93	55
Per Cent.....	100	17.9	1.0	18.8	27.5
Nine to Ten					
Number.....	184	3	3	6	6
Per Cent.....		1.6	1.6	3.3	3.0
Nine to Eleven					
Number.....	121	3	2	5	2
Per Cent.....		2.5	1.7	4.1	1.0

(concluded on next page)

Table 30. (concluded)

Grades Tested	Number of Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Different Students Involved
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Total					
Number.....	305	6	5	11	8
Per Cent.....	100	2.0	1.6	3.6	4.0
Ten to Eleven					
Number.....	113	1	2	3	3
Per Cent.....	100	0.9	1.8	2.7	1.5
Total Test Scores					
Number.....	1563	122	45	167	78
Per Cent.....	100	7.8	2.9	10.7	39

It is interesting to note that the pattern of deviations in the seventh-grade scores differs from that at other levels. Note in column (5) that all percentages of total deviations which exceed 10 per cent involve seventh-grade scores. Observe also in columns (3) and (4) that when the seventh grade is used as a point of reference all but five of the deviations represent plus achievement. When the other grade levels are used as points of reference the deviations are more evenly divided between plus and minus achievement. It may be seen in column (6) that more than half of the students who deviated in reading achievement did so when seventh-grade scores were used. As indicated in column (6), 39 per

cent of the students deviated at least once and 61 per cent worked within a two-stanine range of previous reading achievement when compared with every grade level.

Deviations in arithmetical reasoning.-- It may be observed in Table 31 that the pattern of predominantly plus achievement which existed when seventh-grade reading scores were used as points of reference was also found in arithmetic reasoning. It is interesting to note in columns (3) and (4) that when elementary algebra, which was given in grade nine A, was compared with previous achievement the deviations showed a marked increase in plus achievement; and when compared with subsequent achievement, the deviations were predominantly minus. These findings suggest that students tend to score higher in elementary algebra than in other tests which involve mathematical reasoning. These results might be due to the different standardization samples; however, this trend was not evident among the students who took the intermediate form of the algebra test.

Table 31. A Comparison of the Stanine Variations of Three or More in the Arithmetic Reasoning Test Scores of the Same Group of 200 Students As They Progressed From the Sixth Grade to the Eleventh

Grades Tested	Number of Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Different Students Involved
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Six to Seven					
Number.....	181	7	15	22	22
Per Cent.....		3.9	8.3	12.2	11.0

(continued on next page)

Table 31. (continued)

Grades Tested	Number of Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Different Students Involved
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Six to Nine B					
Number.....	182	9	12	21	16
Per Cent....		4.9	6.6	11.5	8.0
Six to Nine A					
Number.....	153	17	4	21	15
Per Cent....		11.1	2.6	13.7	7.5
Six to Eleven					
Number.....	111	8	3	11	2
Per Cent....		7.2	2.7	9.9	1.0
Total					
Number.....	627	41	34	75	55
Per Cent....	100	6.5	5.4	12.0	27.5
Seven to Nine B					
Number.....	180	13	5	18	18
Per Cent....		7.2	2.8	10.0	9.0
Seven to Nine A					
Number.....	162	29	5	34	27
Per Cent....		17.9	3.1	21.0	13.5
Seven to Eleven					
Number.....	117	15	2	17	4
Per Cent....		12.8	1.7	14.5	2.0
Total					
Number.....	459	57	12	69	49
Per Cent....	100	12.4	2.6	15.0	24.5
Nine B to Nine A					
Number.....	165	25	6	31	31
Per Cent....		15.2	3.6	18.8	15.5
Nine B to Eleven					
Number.....	120	6	2	8	3
Per Cent....		5.0	1.7	6.7	1.5

(concluded on next page)

Table 31. (concluded)

Grades Tested	Number of Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Different Students Involved
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Total					
Number.....	285	31	8	39	34
Per Cent.....	100				17.0
Nine A to Eleven					
Number.....	114	3	16	19	19
Per Cent.....	100	2.6	14.0	16.7	9.5
Total Test Scores					
Number.....	1485	132	70	202	91
Per Cent.....	100	8.8	4.7	13.6	45.5

As may be seen in column (6), approximately 46 per cent of the students showed at least one deviation in arithmetic reasoning, and about 54 per cent worked within a two-stanine range of their previous achievement at all grade levels.

Deviations in English.-- The comparison of achievement scores with previous achievement at different grade levels indicated that the highest percentage of significant deviations occurred in English. Column (3) of Table 32 shows that when the seventh-grade scores were used as points of reference the differential achievement was predominantly plus. This is the same pattern that was found in both reading and arithmetic reasoning. Note also that a similar pattern

existed when the tenth-grade scores were used as points of reference. The identification of the students who were involved in the significant deviations in the first half of the tenth grade showed that approximately 85 per cent of the deviations were made by students who had entered high school in the tenth grade. This is an interesting finding and may indicate adjustment to a new school may interfere with maximum achievement.

It may be seen in column (6) that 57 per cent of the students deviated at least once, and 43 per cent of the students worked within a normal range at all grade levels. A higher percentage of students deviated in English achievement than in any other area.

Table 32. A Comparison of the Stanine Variations of Three or More in the English Scores of the Same Group of 200 Students As They Progressed From the Seventh Grade to the Eleventh

Grades Tested	Number of Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Different Students Involved
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Seven to Nine					
Number.....	183	26	3	29	29
Per Cent.....		14.2	1.6	15.8	14.5
Seven to Ten B					
Number.....	177	7	24	31	26
Per Cent.....		4.0	13.6	17.6	13.0

(continued on next page)

Table 32. (continued)

Grades Tested	Number of Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Different Students Involved
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Seven to Ten A					
Number.....	165	26	13	39	17
Per Cent....		15.8	7.9	23.7	8.5
Seven to Eleven					
Number.....	107	20	3	23	8
Per Cent....		18.7	2.8	21.5	4.0
Total					
Number.....	632	79	43	122	80
Per Cent....	100	12.5	6.8	19.3	40.0
Nine to Ten B					
Number.....	191	8	37	45	45
Per Cent....		4.2	19.4	23.6	22.5
Nine to Ten A					
Number.....	177	8	19	27	14
Per Cent....		4.5	10.7	15.2	7.0
Nine to Eleven					
Number.....	120	8	2	10	5
Per Cent....		6.7	1.7	8.4	2.5
Total					
Number.....	488	24	58	82	64
Per Cent....		4.9	11.9	16.8	32.0
Ten B to Ten A					
Number.....	171	19	10	29	29
Per Cent....		11.1	5.8	16.9	14.5
Ten B to Eleven					
Number.....	114	19	0	19	14
Per Cent....		16.7	0	16.7	7.0

(concluded on next page)

Table 32. (concluded)

Grades Tested	Number of Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Different Students Involved
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Total					
Number.....	285	38	10	48	43
Per Cent.....	100	13.3	3.5	16.8	21.5
Ten A to Eleven					
Number.....	115	14	2	16	0
Per Cent.....	100	12.2	1.7	13.9	0
Total Test Scores					
Number.....	1520	155	113	268	114
Per Cent.....	100	10.2	7.4	17.6	57.0

Deviations in vocabulary.-- A glance at Table 33 shows that all deviations in vocabulary represented plus achievement. Thirty different students were responsible for the deviations when seventh-grade scores were used as points of reference. The two students who deviated between grades nine and eleven had also deviated between grades seven and eleven. This was the only area in which no minus achievement was found between any grade levels.

Table 33. A Comparison of the Stanine Variations of Three or More in the Vocabulary Scores of the Same Group of 200 Students As They Progressed From the Seventh Grade to the Eleventh

Grades Tested	Number of Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Different Students Involved
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Seven to Nine					
Number.....	195	26	0	26	26
Per Cent.....		13.3	0	13.3	13.0
Seven to Eleven					
Number.....	118	12	0	12	4
Per Cent.....		10.2	0	10.2	2.0
Total					
Number.....	313	38	0	38	30
Per Cent.....	100	12.1	0	12.1	15.0
Nine to Eleven					
Number.....	121	2	0	2	2
Per Cent.....	100	1.7	0	1.7	1.0
Total Test Scores					
Number.....	434	40	0	40	30
Per Cent.....	100	9.2	0	9.2	15.0

Deviations in arithmetic computation and spelling.-- In these two areas, only the seventh and ninth-grade scores were available for comparison. Note in columns (3) and (4) of Table 34 that the number of plus and minus deviations in arithmetic computation is approximately the same. This is the only subject in which the pattern of plus achievement

did not prevail when the seventh-grade achievement was used as a point of reference. Nineteen of the twenty-one deviations in spelling were plus. Three students deviated in both arithmetic computation and spelling.

Table 34. A Comparison of the Stanine Variations of Three or More in the Arithmetic Computation and Spelling Scores of the Same Group of 200 Students in Grades Seven and Nine

Subject	Number of Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Different Students Involved
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Arithmetic Computation					
Number.....	180	9	10	19	19
Per Cent....	100	5.0	5.6	10.6	9.5
Spelling					
Number.....	195	19	2	21	21
Per Cent....	100	9.7	1.0	10.7	10.5
Total Test Scores					
Number.....	375	28	12	40	37
Per Cent....	100	7.5	3.2	10.7	18.5

Deviations in all subjects.-- The calculation of all possible stanine differentials between achievement and previous achievement indicated that the greatest percentage of deviations occurred in English. Identification of the different students involved in the deviations from grade to

grade showed that more students deviated in English than in any other subject.

Table 35 shows the percentages of deviations when different grade levels were used as points of reference. Note in column (5) that the highest percentage of deviations occurred when the seventh-grade level was used as a point of reference. The comparatively large percentage of plus deviations at this level indicates, as has been shown throughout this study, that students tend to score lower in all subjects at the seventh-grade level than at any other grade level from six to eleven. Of the 148 students responsible for the deviations when seventh-grade scores were compared with subsequent scores, 98 represented plus achievement in all areas.

As may be seen in column (6), 86 per cent of the students deviated at least once, and 14 per cent functioned within a two-stanine range throughout the five-year period when achievement was compared with previous achievement. Twenty-eight per cent of the students showed only plus achievement, and more than fifty per cent showed two or less deviations in all subjects at all levels.

In general, there is little chance that the percentage of significant deviations between achievement and previous achievement at all levels will be less than 12 or more than 15 per cent.

Table 35. A Comparison of the Stanine Variations of Three or More in the Achievement Scores of the Same Group of 200 Students When Different Grade Levels Are Used As Points of Reference

Point of Reference	Number of Total Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Different Students Involved
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Grade Six					
Number.....	1276	68	67	135	84
Per Cent....	100	5.3	5.2	10.5	42.0
Grade Seven					
Number.....	2275	290	72	362	148
Per Cent....	100	12.7	3.2	15.9	74.0
Grade Nine					
Number.....	1313	66	87	153	94
Per Cent....	100	5.0	6.6	11.6	47.0
Grade Ten					
Number.....	513	53	14	67	35
Per Cent....	100	10.3	2.7	13.0	17.5
Total Test Scores					
Number.....	5377	477	240	717	172
Per Cent....	100	8.9	4.5	13.4	86.0

12. Total Stability in Achievement Test Scores

At all Grade Levels

Stability in reading.-- Four different points of reference were used to determine the total stability in reading between achievement and previous achievement. The stability ranged from approximately 81 to 97 per cent with a

composite stability of approximately 89 per cent. Note in column (6) of Table 36 that the seventh grade was the only level at which the percentage dropped below 90 per cent. As it has been pointed out previously, the deviations from the seventh-grade level showed a marked tendency toward plus achievement. In general a high degree of stability exists between reading scores at different grade levels. Critical ratios of 4.4 or higher indicated that it is almost certain that the percentage of stability between reading scores at all grade levels will be lower when seventh-grade scores are used as points of reference than when other levels are used.

Table 36. A Comparison of the Stanine Stability in the Reading Performance of the Same Group of 200 Students in Grades Six to Eleven When Different Grade Scores Are Used As Points of Reference

Point of Reference	Number of Total Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Grade Six					
Number.....	649	27	33	60	589
Per Cent.....		4.2	5.1	9.2	90.8
Grade Seven					
Number.....	496	88	5	93	403
Per Cent.....		17.9	1.0	18.8	81.2
Grade Nine					
Number.....	305	6	5	11	294
Per Cent.....		2.0	1.6	3.6	96.4
Grade Ten					
Number.....	113	1	2	3	110
Per Cent.....		0.9	1.8	2.7	97.3
Total Test Scores					
Number.....	1563	122	45	167	1396
Per Cent.....	100	7.8	2.9	10.7	89.3

Stability in arithmetic reasoning.-- When different grade levels were used as points of reference, the stanine stability in arithmetic reasoning ranged from 83 to 88 per cent, with a composite stability of approximately 86 per cent. No significant differences were found between the percentages of stability at different levels; but, as shown in column (3) of Table 37, the deviations from both grade seven and nine B achievement were primarily plus.

Table 37. A Comparison of the Stanine Stability in the Test Scores in Arithmetic Reasoning of the Same Group of 200 Students in Grades Six to Eleven When Different Grade Scores Are Used As Points of Reference

Point of Reference	Number of Total Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Grade Six					
Number.....	627	41	34	75	552
Per Cent....		6.5	5.4	12.0	88.0
Grade Seven					
Number.....	459	57	12	69	390
Per Cent....		12.4	2.6	15.0	85.0
Grade Nine B					
Number.....	285	31	8	39	246
Per Cent....		10.9	2.8	13.7	86.3
Grade Nine A					
Number.....	114	3	16	19	95
Per Cent....	100	2.6	14	16.7	83.3
Total Test Scores					
Number.....	1485	132	70	202	1283
Per Cent....	100	8.8	4.7	13.6	86.4

Stability in English.-- Table 38 shows the stability in English scores when different grade levels were used as reference points. It may be seen in Column (6) that the stability varied from approximately 81 to 86 per cent with a composite stability of approximately 82 per cent. No significant differences were found between any grade levels, but observe in columns (3) and (4) that the ninth-grade reference point was the only one at which the majority of deviations represented minus achievement.

Table 38. A Comparison of the Stanine Stability in the English Scores of the Same Group of 200 Students In Grades Seven to Eleven When Different Grade Scores Are Used As Points of Reference

Point of Reference	Number of Total Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Grade Seven					
Number.....	632	79	43	122	510
Per Cent....		12.5	6.8	19.3	80.7
Grade Nine					
Number.....	488	24	58	82	406
Per Cent....		4.9	11.9	16.8	83.2
Grade Ten B					
Number.....	285	38	10	48	237
Per Cent....		13.3	3.5	16.8	83.2
Grade Ten A					
Number.....	115	14	2	16	99
Per Cent....		12.2	1.7	13.9	86.1
Total Test Scores					
Number.....	1520	155	113	268	1252
Per Cent....	100	10.2	7.4	17.6	82.4

Stability in vocabulary.-- Only two reference points were available for comparison of vocabulary scores.

Although a limited number of differentials were computed in this area it is interesting to note in Table 39 that no student scored below the normal stanine range of his previous achievement.

Table 39. A Comparison of the Stanine Stability in the Vocabulary Scores of the Same Group of 200 Students in Grades Seven to Eleven When Different Grade Scores Are Used As Points of Reference

Point of Reference	Number of Total Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Grade Seven					
Number.....	313	38	0	38	275
Per Cent.....		12.1	0	12.1	87.9
Grade Nine					
Number.....	121	2	0	2	119
Per Cent.....		1.7	0	1.7	98.3
Total Test Scores					
Number.....	434	40	0	40	394
Per Cent.....	100	9.2	0	9.2	90.8

In order to give a complete picture of the composite stability when all possible differentials were computed between subject scores at all levels Table 40 was prepared. A glance at column (6) shows that the stability ranged from

Table 40. A Comparison of the Stanine Stability By Subject in the Performance of the Same Group of 200 Students in Grades Six to Eleven When All Possible Differentials Were Computed Between Subject Scores At Various Grade Levels

Name of Subject	Number of Total Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Reading					
Number.....	1563	122	45	167	1396
Per Cent.....		7.8	2.9	10.7	89.3
Arithmetic Reasoning					
Number.....	1485	132	70	202	1283
Per Cent.....		8.8	4.7	13.6	86.4
English Usage					
Number.....	1520	155	113	268	1252
Per Cent.....		10.2	7.4	17.6	82.4
Vocabulary					
Number.....	434	40	0	40	394
Per Cent.....		9.2	0	9.2	90.8
Arithmetic Computation					
Number.....	180	9	10	19	161
Per Cent.....		5.0	5.6	10.6	89.4
Spelling					
Number.....	195	19	2	21	174
Per Cent.....		9.7	1.0	10.7	89.2
Total Test Scores					
Number.....	5377	477	240	717	4660
Per Cent.....	100	8.9	4.5	13.4	86.6

82.4 to 90.8 per cent in each subject area, with a composite stability of approximately 87 per cent for all subjects.

English showed the lowest degree of stability, and vocabulary,

the highest. In general a substantial to high degree of stanine stability exists in every subject area when all possible reference points are used for comparison of achievement with previous achievement. Significant deviations tend to represent plus achievement in all subjects except arithmetic computation.

Composite stability by achievement tests.-- The data in Table 41 show the percentages of stanine stability of the composite scores when different achievement tests were used as points of reference. Note how little the percentages vary both in stanine stability and plus and minus achievement when either the sixth or ninth-grade scores were used as reference points. It is interesting to note also that when seventh-grade scores were the reference points, the per cent of plus achievement was almost triple that of minus achievement. These findings suggest that, although a substantial degree of stability exists between seventh-grade scores and those of other grades, students do not achieve as well in the seventh grade as in other grades. This trend appears to exist at the beginning of senior high school also, as indicated when the tenth-grade scores were used as points of reference.

The results of the calculation of all possible stanine differentials between achievement test scores over a five-year period indicated that there is little possibility that the total stability will drop below 80 per cent in English usage,

Table 41. A Comparison of the Stanine Stability in the Composite Scores of All Achievement Tests of the Same Group of 200 Students in Grades Six to Eleven When Different Tests Are Used As Points of Reference

Names of Ability and Achievement Tests	Number of Total Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Metropolitan Achievement Tests					
Grade Six					
Number.....	1276	68	67	135	1141
Per Cent....	100	5.3	5.2	10.5	89.4
Stanford Achievement Test					
Grade Seven					
Number.....	2275	290	72	362	1913
Per Cent....	100	12.7	3.2	15.9	84.1
Grade Nine					
Number.....	1199	63	71	134	1065
Per Cent....	100	5.3	5.9	11.2	88.8
Total					
Number.....	3474	353	143	496	2978
Per Cent....	100	10.2	4.1	14.3	85.7
Cooperative Tests					
Grade Nine A					
Number.....	114	3	16	19	95
Per Cent....	100	2.6	14.0	16.7	83.3
Grade Ten					
Number.....	513	53	14	67	446
Per Cent....		10.3	2.7	13.1	86.9
Total					
Number.....	627	56	30	86	541
Per Cent....	100	8.9	4.8	13.7	86.3
Total Test Scores					
Number.....	5377	477	240	717	4660
Per Cent....	100	8.9	4.5	13.4	86.6

or below 84 per cent in any other subject area. The highest degrees of stability were found when either the sixth or ninth-grade scores were used as reference points, although a substantial degree of stability existed at other levels.

The difference between the composite stability in reading and arithmetic reasoning, and that in all subjects used in this study was not significant. This finding suggests that the stability in the composite scores of reading and arithmetic reasoning is a fairly reliable indication of the stability that may be expected in the composite scores of all skill subjects.

13. Ability and Previous Achievement as

Reference Points in Stanine Stability

Grade Six to Eleven.-- It will be recalled that the prime purpose of this study was to evaluate the use of sixth-grade national stanine scores in ability, and those in the subtests of reading and arithmetic problems as criteria of potential high-school achievement. In order to compare the percentages of significant deviations obtained when ability was used as a reference point with those obtained when sixth-grade achievement was used as a reference point, Table 42 was devised.

Deviation patterns.-- It is interesting to note that not only were the highest percentages of deviations from both reference points found at the seventh-grade level, but the

Table 42. A Comparison of the Stanine Variations of Three or More in the Performance of the Same Group of 200 Students in Grades Seven to Eleven When Ability Is Used As a Reference Point and When Sixth-Grade Achievement Is Used As a Reference Point

Grade and Direction of Variations	Per Cent of Scores Varying According to Reference Points Indicated					
	Reading		Arithmetic Reasoning		Total Test Scores	
	Abil-ity	Achieve-ment	Abil-ity	Achieve-ment	Abil-ity	Achieve-ment
(1)	(2)		(3)		(4)	
Grade Seven						
Plus....	0.0	2.7	0.5	3.9	0.3	3.3
Minus...	31.0	11.4	22.5	8.3	26.7	9.9
Total...	31.0	14.1	23.0	12.2	27.0	13.2
Grade Nine B						
Plus....	0.5	6.5	0.0	4.9	0.3	5.7
Minus...	8.8	3.3	18.8	6.6	13.7	4.9
Total...	9.3	9.8	18.8	11.5	14.0	10.7
Grade Nine A						
Ten						
Plus....	1.1	3.0	3.0	11.1	2.0	6.8
Minus...	4.4	1.8	7.8	2.6	6.0	2.2
Total...	5.5	4.7	10.8	13.7	8.0	9.0
Grade Eleven						
Plus....	0.0	4.5	0.8	7.2	0.4	5.8
Minus...	5.8	2.7	6.7	2.7	6.2	2.7
Total...	5.8	7.1	7.6	9.9	6.6	8.5
Total Test Scores						
Plus....	0.4	4.2	1.0	6.5	0.7	5.3
Minus...	13.3	5.1	15.0	5.4	14.1	5.2
Total...	13.7	9.3	16.0	11.9	14.8	10.5

deviation patterns at this level differed from those at other levels. Observe that in grade seven (1) the percentages of deviations from ability were approximately double those from achievement in both subjects, and (2) all deviations were primarily minus from both reference points. Note, however, that these patterns did not exist at other levels. In all other grades (1) no significant differences were found between the percentages of deviations from ability and those from achievement, (2) when ability was the reference point the deviations were nearly all minus, and (3) when achievement was the reference point the deviations tended to indicate plus achievement.

It would appear from these data that when sixth-grade records are used as criteria of potential achievement of students with an ability stanine of six or above, more students will score within or above their normal stanine range when achievement is the reference point than when ability is the reference point. There is little chance that the composite stability in reading and arithmetic reasoning over a five-year period will drop below 87 per cent when sixth-grade achievement is used as a point of reference, or below 82 per cent when ability is used as a point of reference. As has been found throughout this study, the seventh-grade scores will show the highest percentage of deviations. If the seventh-grade scores are omitted from the total

stability, the percentage of stability between ability and achievement will increase to a minimum of 87 per cent.

From this data it may be concluded that the sixth-grade achievement record in reading and arithmetic problems is a comparatively reliable indication of how well students will achieve in these two subjects at the high-school level. Since no significant differences were found between the composite stability in these two subjects, and that in all subject areas at other grade levels, it may be assumed that the composite achievement in reading and arithmetic reasoning is a general indication of composite achievement in all skill subjects.

Total stability.-- Table 43 contains the results of the computation of all possible stanine differentials between ability and achievement, achievement and previous achievement, and a composite of both. Although the total stability between ability and achievement, and that between achievement and previous achievement varied less than two percentage points, the difference was significant at the 1 per cent level.

It will be observed in columns (3) and (4) that when ability was the reference point, the percentage of minus deviations greatly exceeded that of plus deviations; and when previous achievement was the reference point, the opposite pattern prevailed. Critical ratios indicated that highly significant differences existed between the percentages of

Table 43. A Comparison of the Stanine Stability in the Composite Scores of the Same Group of 200 Students in Grades Six to Eleven When All Possible Differentials Were Computed Between Scores at All Grade Levels

Points of Reference	Number of Total Test Scores	Number, Per Cent, and Direction of Stanine Variations of Three or More			Number and Per Cent of Variations of Two or Less
		Plus	Minus	Total	
(1)	(2)	(3)	(4)	(5)	(6)
Ability Scores					
Number.....	6212	73	863	936	5276
Per Cent.....	100	1.2	13.9	15.1	84.9
Previous Achievement Scores					
Number.....	5377	477	240	717	4660
Per Cent.....	100	8.9	4.5	13.4	86.6
Total Test Scores					
Number.....	11589	550	1103	1653	9936
Per Cent.....	100	4.7	9.5	14.2	85.7

deviations in the areas of (1) overachievement and underachievement, (2) plus and minus achievement, (3) overachievement and plus achievement, and (4) underachievement and minus achievement. The data in Table 43 show conclusively that there is a far greater chance of plus achievement than there is of overachievement. However, in order to view these results in their true perspective, it must be remembered that one of the limitations of stanine differentials is that any student with an ability stanine of seven or higher can not be identi-

1783
fied as an overachiever unless an 11 point scale is used to separate the 1 per cent at either end of the curve. This fact is especially applicable to this group as 71 per cent of the students in this study had an ability stanine of seven or above. Since the purpose of this study was not to identify overachievers, but to identify potentially superior high-school students, the stanine methodology was considered adequate.

The results of the calculation of 11,589 stanine differentials indicated that it is almost certain that the total stanine stability will not be less than 85 per cent, irrespective of points of reference, when students with above average ability are tested over a five-year period. In general students tend to deviate less from previous achievement than from ability in all areas at all grade levels in grades six to eleven.

14. Sixth-Grade Ability and Achievement As Criteria of High-School Achievement

The preceding findings indicate the range of achievement which may be expected over a five-year period. However, since the purpose of this study was to identify the potentially superior high-school student, it was deemed necessary to determine the actual achievement stanines attained at the high-school level.

Measures of high-school achievement.-- Two measures were used as criteria of general high-school achievement. The first measure, which was available for all 200 students, was the average stanine score on all Cooperative Tests at the ninth and tenth-grade levels. These tests were used as measures of achievement in the following areas:

1. Elementary algebra, elementary languages, reading, and English for all students
2. Intermediate algebra and advanced Latin, or the sciences, according to the subjects elected in high school.

Seven test scores were available for the majority of the students.

The second criterion measure was the composite stanine score on the Iowa Tests of Educational Development which included nine subject areas. As described earlier in this chapter, these tests were administered to 121 eleventh-grade students who had scored in stanine 6 or above on the sixth-grade ability test. One student was unable to complete the entire battery of tests which left a total of 120 composite scores available for comparison with sixth-grade records.

A stanine score of 7 or higher was considered to indicate slightly superior to very superior achievement; a stanine score of 6 was considered to indicate high average achievement.

Four bivariate distributions were plotted to show the relationship and trend of expectancy between the following measures:

1. Sixth-grade ability and Cooperative average stanine scores
2. Sixth-grade ability and ITED composite stanine scores
3. Sixth-grade average achievement and Cooperative average stanine scores
4. Sixth-grade average achievement and ITED composite stanine scores.

Ability stanines as criteria.-- Table 44 shows the bivariate distribution between sixth-grade ability and the Cooperative average stanine scores. Observe that none of the 33 students with an ability stanine of 9 had an average achievement stanine of less than 6, and 25 scored in stanine 7 or above. Roughly speaking, the chances are 25 out of 33 that students with a sixth-grade ability stanine of 9 will show slightly superior to superior achievement at the high-school level. Conversely, the chances are 10 out of 17 that students with an ability stanine below 6 will achieve in the low to below average range. The correlation between these two measures was 0.89, which indicates a high degree of relationship.

Table 44. Bivariate Distribution Showing Relationship Between Sixth-Grade Ability Stanines and High-School Average Achievement Stanines

Sixth-Grade Ability Stanine	Number and Per Cent of Students Attaining Average Achievement Stanine Scores					Number and Per Cent of Total Scores
	Below 5	5.0 to 5.9	6.0 to 6.9	7.0 to 7.9	8.0 to 8.9	
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Stanine 9						
Number.....	0	0	8	13	12	33
Per Cent...	0	0	24.2	39.4	36.4	16.5
Stanine 8						
Number.....	6	16	24	19	5	70
Per Cent...	8.6	22.9	34.4	27.1	7.1	35.0
Stanine 7						
Number.....	12	8	9	7	2	38
Per Cent...	31.6	21.1	23.7	18.4	5.3	19.0
Stanine 6						
Number.....	18	9	13	1	1	42
Per Cent...	42.8	21.4	31.0	2.4	2.4	21.0
Below Stanine 6						
Number.....	10	5	2	0	0	17
Per Cent...	58.8	29.4	11.8	0	0	8.6
Total Scores						
Number.....	46	38	56	40	20	200
Per Cent...	23.0	19.0	28.0	20.0	10.0	100

The relationship between sixth-grade ability and the ITED composite score is shown in Table 45. As was found with the Cooperative Tests, none of the students with an ability stanine of 9 scored below stanine 6 on the ITED composite score; and 84 per cent achieved in stanine 7 or higher. As may be seen in column (1), no student with an ability stanine below 6 was

Table 45. Bivariate Distribution Showing Relationship Between Sixth-Grade Ability Stanines and Composite Stanine Scores on the Iowa Tests of Educational Development

Grade-Six Ability Stanine	Number and Per Cent of Students Scoring in ITED Stanines					Number and Per Cent of Total Scores
	5	6	7	8	9	
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Stanine 9						
Number.....	0	5	9	8	9	31
Per Cent....	0	16.1	29.0	25.8	29.0	25.8
Stanine 8						
Number.....	2	16	17	16	1	52
Per Cent....	3.8	30.8	32.7	30.8	1.9	43.3
Stanine 7						
Number.....	4	12	5	3	0	24
Per Cent....	16.7	50.0	20.8	12.5	0	20.0
Stanine 6						
Number.....	3	3	6	1	0	13
Per Cent....	23.1	23.1	46.2	7.7	0	10.8
Total Scores						
Number.....	9	36	37	28	10	120
Per Cent....	7.5	30.0	30.8	23.3	8.3	100

included in this test. It is interesting to note that less than 8 per cent of the entire group scored below stanine 6, and approximately 63 per cent scored in stanine 7 or higher. The correlation between ability at the sixth-grade level and the ITED at the eleventh-grade level was 0.44, which is a rather low coefficient. However, it indicates a substantial degree of relationship for the range of scores.

Although these bivariate distributions are subject to considerable sampling error, they indicate that the majority

of students with an ability stanine of 8 or 9 at the sixth-grade level may be expected to show high average to superior achievement at the high-school level.

Average achievement stanines as criteria.-- Table 46 shows the bivariate distribution obtained when sixth-grade average achievement stanines on the sub-tests of reading and arithmetic problems of the Metropolitan Achievement Tests were compared with the average achievement stanines on the Cooperative Tests.

Table 46. Bivariate Distribution Showing Relationship Between Sixth-Grade Average Achievement Stanines and High-School Average Achievement Stanines

Sixth-Grade Average Achievement Stanine	Number and Per Cent of Students Attaining Average Achievement Stanine Scores					Number and Per Cent of Total Scores
	Below 5	5.0 to 5.9	6.0 to 6.9	7.0 to 7.9	8.0 to 8.9	
(1)	(2)	(3)	(4)	(5)	(6)	(7)
9.0						
Number.....	0	0	0	0	5	5
Per Cent.....	0	0	0	0	100.0	2.7
8.0 to 8.9						
Number.....	0	0	3	5	10	18
Per Cent.....	0	0	16.7	27.8	55.6	9.8
7.0 to 7.9						
Number.....	4	6	14	17	2	43
Per Cent.....	9.3	14.0	32.6	39.5	4.7	23.4
6.0 to 6.9						
Number.....	5	13	23	11	1	53
Per Cent.....	9.3	24.5	43.4	20.8	1.9	28.8
Below 6.0						
Number.....	35	16	12	2	0	65
Per Cent.....	53.8	24.6	18.5	3.1	0	35.3
Total Scores						
Number.....	44	35	52	35	18	184
Per Cent.....	23.9	19.0	28.3	19.0	9.8	100

As may be observed in columns (5) and (6) all but three of the students with a sixth-grade stanine average of 8 or 9 attained some degree of superiority at the high-school level, while only two of the students with a stanine average below 6 did so. The correlation of 0.69 shows a marked relationship between sixth-grade average achievement and Cooperative average achievement.

The bivariate distribution between sixth-grade achievement and the ITED is shown in Table 47. Note that no student with a sixth-grade achievement stanine of 8 or 9 scored below stanine 7 on the ITED, and approximately 96 per cent of the students who showed high average to superior achievement at the sixth-grade level attained high average to superior achievement at the eleventh-grade level. The correlation between sixth-grade achievement and eleventh-grade achievement was 0.64, which denotes a marked relationship between the two measures.

These findings indicate that in general students whose average achievement stanine is 8 or 9 at the sixth-grade level may be expected to show slightly superior to very superior achievement at the high-school level.

Predicting achievement.--- Once again it must be stressed that bivariate distributions are subject to considerable sampling error. However, the results of these findings may

Table 47. Bivariate Distribution Showing Relationship Between Sixth-Grade Average Achievement Stanine Scores and ITED Composite Stanine Scores

Sixth-Grade Average Achievement Stanine Score (1)	Number and Per Cent of Students Scoring in ITED Stanines					Number and Per Cent of Total Scores (7)
	5 (2)	6 (3)	7 (4)	8 (5)	9 (6)	
9.0 Number.....	0	0	0	2	3	5
Per Cent.....	0	0	0	40.0	60.0	4.5
8.0 to 8.9 Number.....	0	0	2	11	4	17
Per Cent.....	0	0	11.8	64.7	23.5	15.3
7.0 to 7.9 Number.....	3	8	14	5	2	32
Per Cent.....	9.4	25.0	43.8	15.6	6.3	28.8
6.0 to 6.9 Number.....	1	12	16	6	0	35
Per Cent.....	2.9	34.3	45.7	17.1	0	31.5
Below 6.0 Number.....	5	13	4	0	0	22
Per Cent.....	22.7	59.1	18.2	0	0	19.8
Total Scores Number.....	9	33	36	24	9	111
Per Cent.....	8.1	29.7	32.4	21.6	8.1	100

be considered as general indications of potential achievement. Thus, it may be assumed that the majority of students who show superior ability or achievement in the sixth-grade criteria measures will show slightly superior to very superior achievement in high school. As has been found throughout this study, the criterion of sixth-grade achievement appears to be a slightly better predictor of potential achievement than that of ability.

CHAPTER V

CONCLUDING SUMMARY, IMPLICATIONS,
LIMITATIONS, AND RECOMMENDATIONS

1. Stability Between Mental Ability
Test Scores

It will be recalled that the mental ability tests used in this study were the California Short-Form Test of Mental Maturity at the sixth-grade level, and the Otis Quick-Scoring Mental Ability Test at the ninth-grade level. When differentials were computed between the national stanine scores on these two tests, 32 per cent of the students were found to have the same stanine score in both tests; approximately 48 per cent varied one stanine; about 15 per cent varied two stanines; and 5 per cent varied three or more stanines. In other words, approximately 80 per cent of the students functioned within a one stanine range, and a total of 95 per cent showed stanine stability in mental ability from grade six to grade nine.

These findings suggest that not only does mental ability show stability between grades six and nine, but it is almost certain that the percentage of stability between the national stanine scores on the two different tests will not drop below 90 per cent.

2. Stability Between Mental Ability and Achievement Test Scores

CTMM-achievement differentials.-- The ability scores on the CTMM were projected through the ninth grade, and compared with the test scores of the Metropolitan Achievement Tests at grade six, and the Stanford Achievement Test at grades seven and nine.

At the sixth-grade level, a stability of approximately 84 per cent was found in both reading and arithmetic problems. As it has been mentioned earlier in this study, only two subtests were used at the sixth-grade level.

At the seventh-grade level, the stability was considerably lower in all six subjects with a range of 69 to 80 per cent, and a total stability of 74 per cent. The highest degree of stability at this level was in word meaning, and the lowest was in reading. This was an interesting finding as both reading and vocabulary are usually considered to be closely related to scholastic aptitude. No significant difference was found between the percentage of stability in the composite scores in reading and arithmetic reasoning, and that in the composite scores in all six subjects at the seventh-grade level.

At the ninth-grade level, a substantial to high degree of stability existed in all subjects with a range of approximately 79 to 95 per cent, and a total stability of 85 per cent.

Word meaning and reading showed the highest degree of stability and spelling showed the lowest. As in the seventh grade, no significant difference was found between the percentage of stability in the composite scores of reading and arithmetic reasoning, and that in the composite scores of all six subjects at the ninth-grade level.

The difference between the total stability in grade six and that in grade nine was not significant. However, the total stability in grade seven was significantly lower than that of either grade six or grade nine.

Otis-achievement differentials.-- The ability scores on the Otis test were projected through grade eleven, and compared with the Stanford Achievement Test at grade nine B, the Cooperative Tests at grades nine A and ten, and the Iowa Tests of Educational Development at grade eleven.

Although ninth-grade achievement scores had already been compared with CTMM ability, a second comparison was made with Otis ability to determine if any significant difference existed between the percentages of stability when different ability tests were used as points of reference. The pattern of stability was the same from both reference points with no significant differences in the percentages of stability in reading, word meaning, arithmetic computation, and spelling. However, a significantly higher stability was found in arithmetic reasoning and language usage when Otis ability was

the reference point. The reason for this difference in only two areas warrants further investigation which was not within the scope of this study.

At the ninth and tenth-grade levels a substantial to high degree of stability existed in all areas of the Cooperative Tests except advanced Latin. Reading, elementary Latin, elementary and intermediate algebra, and European history showed a stability of 90 per cent or higher; English, biology, and elementary French ranged from 80 to 88 per cent; and advanced Latin was approximately 64 per cent. The percentage of stability in advanced Latin was significantly lower than that of elementary Latin, and this finding should be investigated to determine if the Providence course of study is consistent with the information covered in the Cooperative Advanced Latin Test.

At the eleventh-grade level, the percentages of stability reached approximately 90 per cent or higher in all tests except basic social concepts which was 83.5, and natural science background which was 64.5 per cent. The significantly low stability in natural science background suggests that the Providence course of study may be lacking in the sciences prior to the eleventh grade.

Overachievement and underachievement.-- As it has been stated previously, one of the limitations of the stanine methodology is that the student who has an ability stanine

of seven or higher can not be identified as an overachiever. Since only 29 per cent of the students in this study were below stanine seven in ability, there was little opportunity for overachievement. Of the 6,212 differentials computed between ability and achievement test scores, 1.2 per cent indicated overachievement, and 13.9 per cent indicated underachievement. Vocabulary showed the lowest percentages of deviations, and the highest percentages of overachievement at all grade levels.

Subject for subject, the percentage of reading deviations in the seventh grade showed the greatest inconsistency when compared with percentages at other grade levels as the students progressed from the sixth grade to the eleventh. The percentage of deviations in grade seven was twice that in grade six, over three times that in grade nine, and five times that in grades ten and eleven. In all other subject areas the percentages of deviations showed a sharp rise at the seventh-grade level but were comparatively consistent at other grade levels. The highest percentage of different students who deviated in any one subject throughout the study was found in English. This was the only area in which a large number showed a first deviation at the tenth-grade level. The highest percentage of total deviations occurred at the seventh-grade level. It was at this level, also, that the greatest number of students showed their first, and for many

their only, deviation.

This significant increase in underachievement at the seventh-grade level warrants further study, but some possible explanations may be (1) adjustment to new school, students, and teachers; (2) adjustment to departmentalized work; and (3) use of machine-scored answer sheets for the first time.

Total stability.-- Vocabulary showed the highest percentage of overall stability. Approximately 90 per cent of the test scores fell within a two-stanine range in this area in grades seven to eleven. In general, vocabulary may be considered as a comparatively steady index of ability. The total stability in reading, arithmetic reasoning, and English ranged from 85 to 87 per cent. No significant difference was found between the percentage of stability in the composite scores of reading and arithmetic reasoning, and that in the composite scores of all subjects at any one grade level or in all five grades combined.

Grade for grade, the only significant differences between the percentages of total composite stability existed at the seventh-grade level. The composite stability at the seventh-grade level was significantly lower than that found at any of the other four levels.

Implications.-- The sixth-grade record of stanine stability between ability scores and the composite scores of reading and arithmetic problems is a comparatively reliable indication

of the total stability which will be found not only in these two subjects, but in all subject areas in high school. This may have implications for the use of the stanine methodology in the identification of underachievers in the two criteria subjects in the sixth grade, and the inauguration of a remedial program at this level. If the percentage of underachievers can be reduced at this level, based on the findings of this study, there should be a consistently high level of achievement in all subjects as the students progress through high school.

The seventh-grade record is not the best criterion of potential achievement as there is a marked tendency toward underachievement in all subjects at this level. An effective remedial program in the skill subjects for underachievers in the seventh grade should greatly improve the chances of a high degree of stability in all subsequent achievement, for a relatively small percentage of students show a first deviation after the seventh-grade level.

3. Stability Between Achievement

Test Scores

Four different achievement test batteries were used in this study: (1) the Metropolitan Achievement Tests in grade six; (2) the Stanford Achievement Test in grades seven and nine; (3) the Cooperative Tests in grades nine A and ten; and

(4) the Iowa Tests of Educational Development in grade eleven. All possible stanine differentials were computed between scores on the sub-tests of each test battery at various grade levels.

Grade six to eleven.-- As it has been stated many times throughout this study, only the two sub-tests of reading and arithmetic problems were used at the sixth-grade level.

When the sixth-grade scores were the reference points a stability of 90 per cent or higher existed in reading at all grade levels except the seventh which showed a stability of approximately 86 per cent. In arithmetic reasoning, the stability ranged from 86 to 90 per cent with no significant differences between any grade levels.

Grade seven to eleven.-- The scores on all six sub-tests were used as reference points for the comparison of seventh-grade achievement with all subsequent achievement. In the two criteria areas the stability was significantly lower at all grade levels, except in arithmetic reasoning in grade nine B, than that found when the sixth-grade scores were the reference points. Stability ranged from 76 to 84 per cent in English, and 87 to 90 per cent in word meaning, arithmetic computation, and spelling.

Grade nine to eleven.-- A very high degree of stability, approximately 96 per cent, existed in reading when grade nine was the reference point. In arithmetic reasoning, the stability dropped to 81 per cent in nine A, but increased to 93 per

cent in grade eleven. The greatest grade to grade variations were found in English. The stability in this area dropped to 76 per cent in ten B, increased to 85 per cent in ten A, and to 91 per cent in grade eleven. Although a limited number of scores were available in vocabulary, 98 per cent of the students indicated stability in grade eleven.

Plus and minus achievement.-- Vocabulary showed a most interesting pattern of deviations. Approximately 9 per cent of the test scores, and 15 per cent of the students showed significant deviations from grades seven to eleven. All deviations indicated plus achievement.

In the subject areas where approximately 1500 stanine differentials were computed, reading showed the lowest percentage of deviations both in the number of scores and in the number of different students involved in the deviations. English showed the highest percentage in both comparisons. The patterns of plus and minus achievement were similar in all subjects except English. The general pattern which was evident throughout this longitudinal study indicated a slight tendency toward plus achievement when previous achievement was used as a reference point at all grade levels except the seventh. At this level there was a marked increase in minus achievement when compared with sixth-grade scores, and a sharp rise in plus achievement when the seventh-grade scores were used as reference points for all subsequent achievement.

In English, the pattern of high plus achievement from the seventh-grade scores prevailed at all levels except ten B where the opposite pattern was found. Significant deviations from ten B English indicated minus achievement when compared with previous achievement, and plus achievement when used as a reference point for subsequent achievement.

It is evident from these findings that students tend to score at a significantly lower achievement level in all subjects in the seventh grade. Similar findings were reported by Corcoran^{1/} in his study on the use of local stanines.

The tendency toward lower achievement exists in English in grade ten B, which suggests additional evidence that adjustment to a new school may interfere temporarily with maximum achievement. This possibility is further strengthened by the fact that the majority of the students who deviated in ten B English were those who had recently enrolled in senior high school, while those who showed stability had been in the same school for a year. This significant difference in stability between the high schools did not exist at subsequent grade levels.

The percentage of the total number of differentials which indicated plus achievement in all subjects was almost

1/John E. Corcoran, A Longitudinal Study to Determine the Stanine Stability of a Group's Test-Score Performance in the Elementary School, Unpublished Doctoral Dissertation, School of Education, Boston University, 1958.

double that which indicated minus achievement.

Total stability.-- When all possible differentials were computed between subject scores at various grade levels, vocabulary, reading, arithmetic computation, and spelling showed a stability range of approximately 89 to 91 per cent; arithmetic reasoning was 86 per cent; and English, 82 per cent. When different grade levels were used as reference points, the composite stability ranged from 84 to 89 per cent. The highest composite stability was found when sixth-grade scores were used as reference points; and the lowest, when seventh-grade scores were the reference points. No significant difference was found between the percentage of stability in the composite achievement in reading and arithmetic reasoning, and that of the composite achievement in all subjects at the same grade level.

Implications.-- All evidence appears to conclude that the sixth-grade composite achievement in reading and arithmetic problems is a good index of potential high-school achievement in all skill areas. Although a substantial degree of stability exists when seventh-grade scores are used as reference points, the general level of achievement tends to be lower in all subjects at the seventh-grade level than at any other grade level. More students deviate significantly when seventh-grade scores are used as reference points than when other grade scores are used.

This lowered achievement at the seventh-grade level may imply that students need greater assistance from both the counselor and subject teachers, in order to adjust more easily during the transitional period. As it was suggested in reference to differential achievement, a remedial program in the skill subjects at the seventh-grade level should greatly improve the chances of a high level of achievement as the students progress through school.

4. Ability and Previous Achievement as Reference Points

When all possible stanine differentials, (6,212) were computed between ability and achievement test scores, the total stability over the five-year period was approximately 85 per cent. When all possible stanine differentials (5,377) were computed between achievement test scores, the five-year stability was approximately 87 per cent. Thus, it may be assumed that, although it is almost certain that the percentage of stability in the total ability-achievement differentials will be less than that of the stability in the total achievement-achievement differentials, the actual difference in percentage points will not be large.

When the percentages of plus and minus deviations in the two types of differentials were compared, the differences in the patterns were striking. The deviations from ability

represented primarily underachievement, while the deviations from previous achievement indicated that the percentage of plus achievement was almost double that of minus achievement.

Approximately the same number of different students deviated from previous achievement as from ability. However, about 33 per cent of the students who deviated from previous achievement represented all plus deviations, while only 2 per cent of those who deviated from ability represented all plus deviations.

The results of the calculation of 11,589 differentials between national stanine scores on different tests indicated that when above-average students are tested it is almost certain that (1) the total stanine stability will not drop below 85 per cent during the five-year period from grades six to eleven, irrespective of points of reference used; and (2) a greater percentage of students will achieve within or above their normal stanine range when previous achievement is used as a reference point than when ability is the reference point.

5. Sixth-Grade Ability and Achievement Stanines

As Criteria of High-School Achievement

The sixth-grade stanine scores were evaluated as criteria of high-school achievement in two ways: (1) as reference points in stanine stability, and (2) as reference points in the prediction of superior high-school achievement.

Prediction of stability.-- The stanine stability in the two criteria subjects of reading and arithmetic problems ranged from 90 to 93 per cent at the high-school level when either sixth-grade ability or achievement scores were used as points of reference. No significant difference was found between the percentage of composite stability in the two criteria subjects and that in all subject areas at any grade level. Therefore, it may be assumed that there is little chance that the percentage of the stanine stability between the sixth-grade ability or composite achievement and high-school composite achievement will be less than 87 per cent. This represents a high degree of stability at the 1 per cent confidence level.

Prediction of superior achievement.-- Stanine stability indicates the range of achievement which may be expected at the high-school level. However, unless the criteria scores are in stanine 9 it does not necessarily identify the potentially superior students. For example, achievement scores in stanines 5 and 6 indicate stability for students with criteria stanine scores of 7 or 8, but they do not denote superior achievement. Although a tighter stanine consistency with differential limits of zero or one narrows the range of expected achievement, a clearer indication of potential achievement was obtained through the use of expectancy tables. The average stanine scores on all Cooperative Tests and the

composite stanine scores on the Iowa Tests of Educational Development were used as measures of high-school achievement. A stanine score of 7 or higher was considered to indicate slightly superior to very superior achievement; a stanine score of 6 was considered to indicate high average achievement. Bivariate distributions between the high-school measures and the criteria measures denoted the following:

1. The majority of students with a sixth-grade ability stanine of 8 or 9 may be expected to show high average to very superior achievement at the high-school level.
2. In general, students with a sixth-grade average achievement stanine of 8 or 9 may be expected to show slightly superior to very superior achievement at the high-school level.
3. The chances are about 7 in 100 that students with a sixth-grade average achievement stanine below 6 will show superior achievement at the high-school level.

Although bivariate distributions are subject to considerable sampling error, they may be considered as general indications of potential achievement. Thus, it may be expected, based on the findings of this study, that the majority of students who show superior ability or achievement in the sixth-grade criteria measures will show slightly superior to very superior achievement at the high-school level.

6. Evaluation of the Stanine Methodology

Advantages.-- Some of the advantages of expressing test scores in stanines are:

1. It permits analysis of mass data on test scores with a minimum of work.
2. Each level represents an equal proportion of the normal curve which permits direct comparison between sub-tests and shows strengths and weaknesses.
3. It permits students to be compared with other students of same age and grade both locally and nationally.
4. It permits measurement of student's growth from year to year.
5. It permits easy identification of superior, average, and below average students.
6. It permits evaluation of achievement in relation to ability.
7. It permits statistical manipulation which may be computed by the average classroom teacher with no training in statistics.
8. It permits easy correlation of two variables.
9. It provides a convenient and uncomplicated technique for deriving valid and meaningful composites.
10. It permits simple and meaningful interpretation of scores to parents and students.
11. Stanines may be directly converted into teachers' marks or grades.

Disadvantages.-- Some of the disadvantages of expressing test scores in stanines are:

1. It does not permit the identification of the 1 per cent of the students at either end of the curve unless an 11-point scale is used.
2. It lacks the refinement of the small unit measures.
3. It does not permit the identification of the superior student who overachieves or the below-average student who underachieves.
4. The difference of 1 percentile point may place a student in a different stanine level.

7. Limitations of the Study

Some of the limitations of this study were:

1. The data consisted of scores on two different ability tests and four different achievement test batteries, each of which was standardized on different, although similar, population groups.
2. The majority of the students tested were above average in ability.
3. Students from only one school system were included.
4. Most of the tests were limited to the skill areas and included little in the content areas.

8. Recommendations

For the city of Providence.-- The results of this study

would seem to indicate the following suggestions for the city of Providence:

1. Consider the possibility of reporting test results in terms of stanine scores.
2. Consider the use of stanine differentials in the identification of underachieving students at the sixth-grade level.
3. Institute an effective remedial program in the skill areas at the sixth-grade level for all students who are achieving below their ability range.
4. Consider classifying junior-high school students according to their stanine scores in the three criteria areas of ability, reading, and arithmetic problems.
5. Institute a strong remedial program in the basic skills and study skills at the seventh-grade level to help underachieving students during the transitional period.
6. Continue to expand the breadth of the work required of junior-high school students with ability stanines of 7 or above in order that they may be sufficiently challenged.

For further study.-- The following are offered as suggestions for further study:

1. Establish local stanine norms and compare them with

national norms to determine whether any significant differences exist between the scores on the two types of norms.

2. Make a similar study using local norms and compare the findings with those of this study to determine whether any significant differences exist at any level or in any subject between the percentages of stanine stability obtained when national norms are used as points of reference and when local norms are used.
3. Make a similar study including a random sampling of all students enrolled in the school system to determine whether the results of this study are applicable to all students or are indicative of above-average students only.
4. Do a follow-up study of the 200 students included in this study to determine (1) how many students accomplished their potential as indicated by their sixth-grade record, and (2) what factors contributed to significant deviations from indicated potential.
5. Make a similar study of stanine stability following the inauguration of a strong remedial program for underachieving students upon entrance into junior-high school to determine whether improvement of basic skills at this level will effect a significantly higher degree of stanine stability between ability and achievement as the student progresses through high school.

APPENDIX

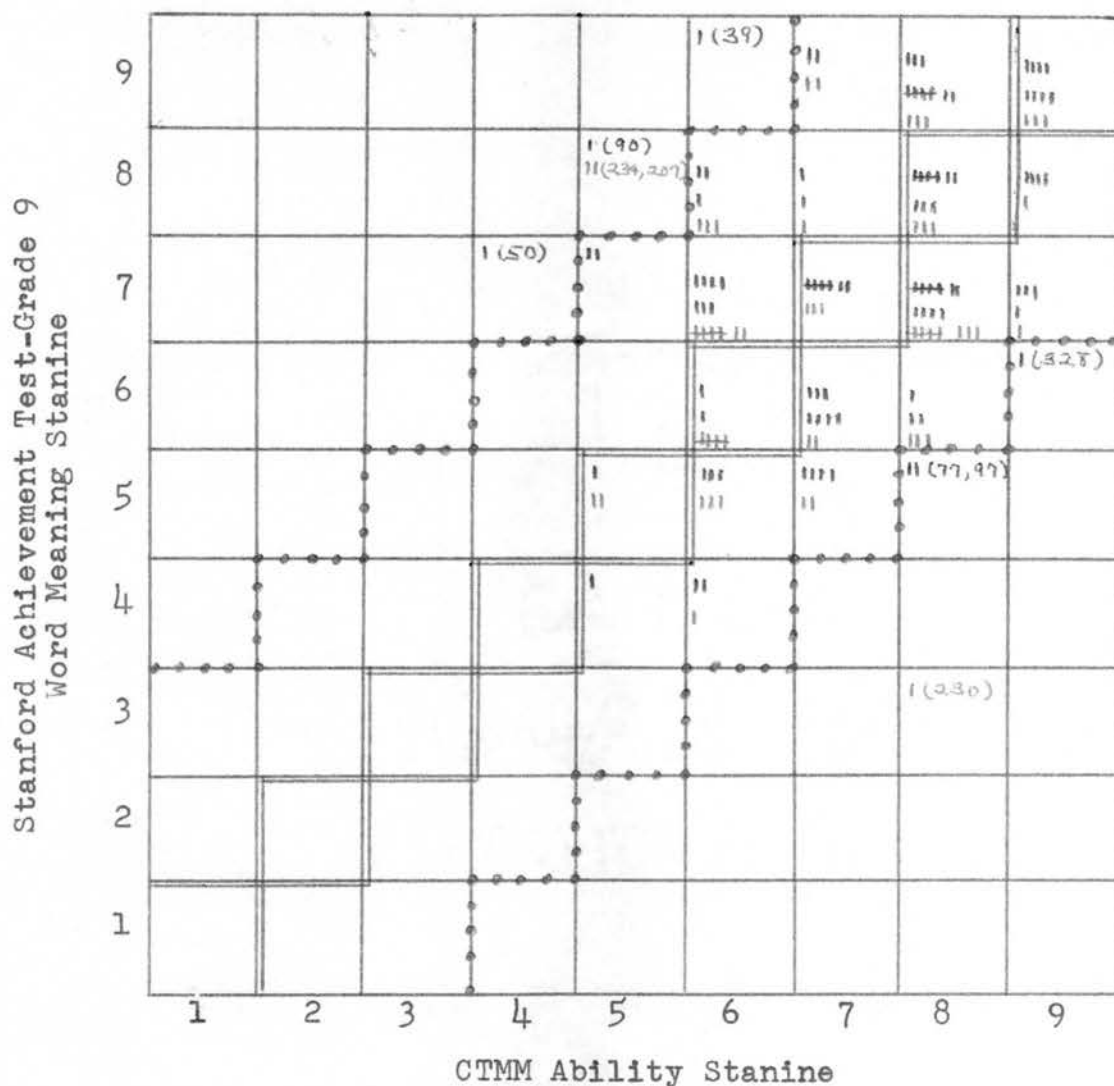


Figure 2. Stanine Chart Showing Identification of Students Whose Achievement Scores Deviated Significantly From Their Ability Scores

The tallies which fall outside of the dotted lines indicate significant deviations or differentials of three or more stanines. The numbers identify the individual student and the colors indicate the group and school from which the students come.

Classical Criterion Hope Criterion
Classical Matched Hope Matched

Table 48. A Comparison of the Stanine Variations in the Scores of the Same Group of 194 Students Between The California Test of Mental Maturity and All Achievement Tests in Grades Six, Seven, and Nine

Grade and Name of Achievement Test	Number of Total Test Scores	Number and Per Cent of Stanine Variations				
		0	1	2	3 or More	Total 0,1,2
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Metropolitan Grade 6						
Number.....	366	97	120	91	58	308
Per Cent.....	100	26.5	32.8	24.9	15.8	84.2
Stanford Grade 7						
Number.....	1145	229	346	274	296	849
Per Cent.....	100	20.0	30.2	23.9	25.9	74.1
Stanford Grade 9						
Number.....	1154	247	458	274	175	979
Per Cent.....	100	21.4	39.7	23.7	15.2	84.8
Total Test Scores						
Number.....	2665	573	924	639	529	2136
Per Cent.....	100	21.5	34.7	24.0	19.8	80.2

Table 49. A Comparison of the Stanine Variations in the Scores of the Same Group of 200 Students Between the Otis Test of Mental Ability and All Achievement Tests in Grades Nine, Ten, and Eleven

Grade and Name of Achievement Test	Number of Total Test Scores	Number and Per Cent of Stanine Variations				
		0	1	2	3 or More	Total 0,1,2
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Stanford Grade 9						
Number.....	1190	351	451	256	132	1058
Per Cent.....	100	29.5	37.9	21.5	11.1	88.9
Cooperative Grades 9A,10						
Number.....	1152	313	405	278	156	996
Per Cent.....	100	27.2	35.2	24.1	13.5	86.5
Iowa Tests of Educational Development Grade 11						
Number.....	1205	358	502	226	119	1086
Per Cent.....	100	29.7	41.7	18.7	9.9	90.1
Total Test Scores						
Number.....	3547	1022	1358	760	407	3140
Per Cent.....	100	28.8	38.3	21.4	11.5	88.5

Table 50. A Comparison of the Stanine Variations in All Ability-Achievement Scores of the Same 200 Students As They Progressed From the Sixth Grade to the Eleventh

Point of Reference	Number of Total Test Scores	Number and Per Cent of Stanine Variations				
		0	1	2	3 or More	Total 0,1,2
(1)	(2)	(3)	(4)	(5)	(6)	(7)
California Test of Mental Maturity						
Number.....	2665	573	924	639	529	2136
Per Cent....	100	21.5	34.7	24.0	19.8	80.2
Otis Test of Mental Ability						
Number.....	3547	1022	1358	760	407	3140
Per Cent....	100	28.8	38.3	21.4	11.5	88.5
Total Test Scores						
Number.....	6216	1595	2282	1399	936	5276
Per Cent....	100	25.7	36.7	22.5	15.1	84.9

Table 51. A Comparison of the Stanine Variations in the Reading Performance of the Same Group of 200 Students in Grades Six to Eleven When Different Grade Scores Are Used As Points of Reference

Point of Reference	Number of Total Test Scores	Number and Per Cent of Stanine Variations				
		0	1	2	3 or More	Total 0,1,2
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Grade Six						
Number.....	649	172	284	133	60	589
Per Cent.....	100	26.5	43.8	20.5	9.2	90.8
Grade Seven						
Number.....	496	108	173	122	93	403
Per Cent.....	100	21.8	34.8	24.6	18.8	81.2
Grade Nine						
Number.....	305	90	152	52	11	294
Per Cent.....	100	29.5	49.8	17.0	3.6	96.4
Grade Ten						
Number.....	113	53	42	15	3	110
Per Cent.....	100	46.9	37.1	13.3	2.7	97.3
Total Test Scores						
Number.....	1563	423	651	322	167	1396
Per Cent.....	100	27.1	41.6	20.6	10.7	89.3

Table 52. A Comparison of the Stanine Variations in the Test Scores in Arithmetic Reasoning of the Same Group of 200 Students in Grades Six to Eleven When Different Grade Scores Are Used As Points of Reference

Point of Reference	Number of Total Test Scores	Number and Per Cent of Stanine Variations				
		0	1	2	3 or More	Total 0,1,2
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Grade Six						
Number.....	627	172	234	146	75	552
Per Cent.....	100	27.4	37.3	23.3	12.0	88.0
Grade Seven						
Number.....	459	104	177	109	69	390
Per Cent.....	100	22.7	38.6	23.7	15.0	85.0
Grade Nine B						
Number.....	285	72	108	66	39	246
Per Cent.....	100	25.3	37.9	23.2	13.7	86.3
Grade Nine A						
Number.....	114	33	39	23	19	95
Per Cent.....	100	28.9	34.2	20.2	16.7	83.3
Total Test Scores						
Number.....	1485	381	558	344	202	1283
Per Cent.....	100	25.7	37.6	23.1	13.6	86.4

Table 53. A Comparison of the Stanine Variations in the Composite Scores of Reading and Arithmetic Reasoning of the Same Group of 200 Students in Grades Six to Eleven When Different Grade Scores Are Used As Points of Reference

Point of Reference	Number of Total Test Scores	Number and Per Cent of Stanine Variations				
		0	1	2	3 or More	Total 0,1,2
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Grade Six						
Number.....	1276	344	518	279	135	1141
Per Cent.....	100	27.0	40.6	21.9	10.6	89.4
Grade Seven						
Number.....	955	212	350	231	162	793
Per Cent.....	100	22.2	36.6	24.2	17.0	83.0
Grade Nine B						
Number.....	590	162	260	118	50	540
Per Cent.....	100	27.4	44.1	20.0	8.5	91.5
Grade Nine A and Ten						
Number.....	227	86	81	38	22	205
Per Cent.....	100	37.9	35.7	16.7	9.7	90.3
Total Test Scores						
Number.....	3048	804	1209	666	369	2679
Per Cent.....	100	26.4	39.7	21.8	12.1	87.9

Table 54. A Comparison of the Stanine Variations By Subject in the Performance of the Same Group of 200 Students in Grades Six to Eleven When All Possible Differentials Were Computed Between Subject Scores At Various Grade Levels

Name of Subject	Number of Total Test Scores	Number and Per Cent of Stanine Variations				
		0	1	2	3 or More	Total 0,1,2
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Reading						
Number.....	1563	423	651	322	167	1396
Per Cent.....	100	27.1	41.6	20.6	10.7	89.3
Arithmetic Reasoning						
Number.....	1485	381	558	344	202	1283
Per Cent.....	100	25.7	37.6	23.1	13.6	86.4
English Usage						
Number.....	1520	362	544	346	268	1252
Per Cent.....	100	23.8	35.8	22.8	17.6	82.4
Vocabulary						
Number.....	434	131	201	62	40	394
Per Cent.....	100	30.2	46.3	14.3	9.2	90.8
Arithmetic Computation						
Number.....	180	48	84	29	19	161
Per Cent.....	100	26.7	46.7	16.1	10.6	89.4
Spelling						
Number.....	195	51	78	45	21	174
Per Cent.....	100	26.2	40.0	23.1	10.8	89.2
Total Test Scores						
Number.....	5377	1396	2116	1148	717	4660
Per Cent.....	100	26.0	39.3	21.3	13.4	86.6

Table 55. A Comparison of the Stanine Variations in the Composite Scores of All Achievement Tests of the Same Group of 200 Students in Grades Six to Eleven When Different Tests Are Used As Points of Reference

Points of Reference	Number of Total Test Scores	Number and Per Cent of Stanine Variations				
		0	1	2	3 or More	Total 0,1,2
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Metropolitan Achievement Tests Grade Six						
Number.....	1276	344	518	279	135	1141
Per Cent.....	100	27.0	40.6	21.9	10.6	89.4
Stanford Achievement Test Grade Seven						
Number.....	2275	531	872	510	362	1913
Per Cent.....	100	23.3	38.3	22.4	15.9	84.1
Grade Nine B						
Number.....	1199	320	498	247	134	1065
Per Cent.....	100	26.7	41.5	20.6	11.2	88.8
Total						
Number.....	3474	851	1370	757	496	2978
Per Cent.....	100	24.5	39.4	21.8	14.3	85.7
Cooperative Tests Grades Nine A and Ten						
Number.....	627	201	228	112	86	541
Per Cent.....	100	32.0	36.4	17.9	13.7	86.3
Total Test Scores						
Number.....	5377	1396	2116	1148	717	4660
Per Cent.....	100	26.0	39.3	21.3	13.4	86.6

Table 56. A Comparison of the Stanine Variations in the Composite Scores of the Same Group of 200 Students in Grades Six to Eleven When All Possible Differentials Were Computed Between Scores At All Grade Levels

Points of Reference	Number of Total Test Scores	Number and Per Cent of Stanine Variations				
		0	1	2	3 or More	Total 0,1,2
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ability Scores						
Number.....	6212	1595	2282	1399	936	5276
Per Cent.....	100	25.7	36.7	22.5	15.1	84.9
Previous Achievement Scores						
Number.....	5377	1396	2116	1148	717	4660
Per Cent.....	100	26.0	39.3	21.3	13.4	86.6
Total Test Scores						
Number.....	11589	2991	4398	2547	1653	9936
Per Cent.....	100	25.8	38.0	22.0	14.2	85.7

Table 57. A Summary of the Total Achievement Stanine Scores of the Same Group of 200 Students As They Progressed From Grade Six to Grade Eleven

Grade and Name of Achievement Test	Number of Total Test Scores	Number and Per Cent of Stanine Scores				
		9	8	7	6	Below 6
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Metropolitan Grade 6						
Number.....	387	15	76	47	120	109
Per Cent...	100	4.1	21.0	12.8	32.7	29.7
Stanford Grade 7						
Number.....	1153	78	124	220	247	484
Per Cent...	100	6.7	10.8	19.1	21.4	42.0
Stanford Grade 9						
Number.....	1190	145	182	283	233	347
Per Cent...	100	12.2	15.3	23.8	19.6	29.1
Cooperative Grades 9A,10						
Number.....	1152	145	207	207	204	389
Per Cent...	100	12.6	18.0	18.0	17.7	33.7
Iowa Tests of Educational Development Grade 11						
Number.....	1205	141	299	303	241	221
Per Cent...	100	11.7	24.8	25.2	20.0	18.3
Total Test Scores						
Number.....	5067	524	888	1060	1045	1550
Per Cent...	100	10.3	17.5	20.9	20.6	30.6

SELECTED BIBLIOGRAPHY

SELECTED BIBLIOGRAPHY

1. Allen, Richard D., et al., Metropolitan Achievement Tests (1947 Edition), Intermediate Battery-Partial. World Book Company, New York.
2. Anastasi, Anne, Psychological Testing, pp. 72-92. The Macmillan Company, New York, 1955.
3. Bixler, Harold H., et al., Metropolitan Achievement Tests, A Guide For Examining a Test Series. World Book Company, New York.
4. Brown, Winnifredian, A Stanine Method of Making Ability-Achievement Comparisons, Unpublished Master's Thesis. Boston University School of Education, Boston, 1958.
5. Buros, Oscar K. (Editor), The Third Mental Measurements Yearbook. Rutgers University Press, New Brunswick, New Jersey, 1949.
6. Coleman, James C. (Editor), California Test of Mental Maturity, Summary of Investigations Number Three. California Test Bureau, Los Angeles, 1956.
7. Conant, James B. (Chairman), The Identification and Education of the Academically Talented Student in the American Secondary School, Invitational Conference on the Academically Talented, 1958. National Education Association, Washington, D. C.
8. Conrad, H. S., "Comparable Measures", pp. 279-282, Encyclopedia of Educational Research, W. S. Monroe, Editor. The Macmillan Company, New York, 1950.
9. Corcoran, John E., A Longitudinal Study to Determine the Stanine Stability of a Group's Test-Score Performance in the Elementary School, Unpublished Doctoral Dissertation. Boston University School of Education, Boston, 1958.
10. Durost, Walter N., The Characteristics, Use, and Computation of Stanines, Test Service Notebook, Number 23. World Book Company, New York, 1959.

11. Durost, Walter N. (Editor), Directions for Administering, Metropolitan Achievement Tests. World Book Company, New York, 1959.
12. Durost, Walter N., and William C. Kvaraceus, The Measurement Notebook. Boston University School of Education, Boston, 1950.
13. Ebel, Robert L., Using Tests for Evaluation, Test Service Notebook, Number 15. World Book Company, New York, 1956.
14. Eckerson, Louise O., "Testing and Counseling", School Life (September, 1960), 43:27-30.
15. Educational Testing Service, Cooperative Tests. 20 Nassau Street, Princeton, New Jersey, 1951.
16. Flanagan, John C., (Editor), The Aviation Psychology Program in the Army Air Forces, Report Number 1, Army Air Forces Aviation Psychology Program Research Reports. United States Government Printing Office, Washington, D. C., 1948.
17. Flanagan, John C., "Units, Scores, and Norms," Chapter XVII in Educational Measurement, E. F. Lindquist, Editor. American Council on Education, Washington, D. C., 1951.
18. Froehlich, Clifford P., and John G. Darley, Studying Students. Science Research Associates, Inc., Chicago, 1952.
19. Guilford, J. P., Fundamental Statistic in Psychology and Education (Second Edition), pp. 303-307. McGraw-Hill Company, New York, 1950.
20. Hart, Irene, Using Stanines to Obtain Composite Scores Based On Test Data and Teachers' Ranks, Test Service Bulletin, Number 86. World Book Company, New York, 1957.
21. Hildreth, Gertrude H., Manual for Interpreting Metropolitan Achievement Tests. World Book Company, New York, 1948.
22. Holmes, John L., Guidance Testing and the Identification of Pupil Characteristics, Educational Bulletin Number 21. California Test Bureau, Del Monte Research Park, Monterey, California, 1959.

23. Jacobs, James N., Aptitude and Achievement Measures in Predicting High School Academic Success, Test Service Bulletin, Number 94. World Book Company, New York, 1959.
24. Johnson, Bertha W., A Study of the Research Done on the Gifted Child from 1952 to the Present Day, Unpublished Master's Thesis, Boston University School of Education, Boston, 1957.
25. Kelley, Truman, et al., Stanford Achievement Test, Advanced Battery-Partial. World Book Company, New York, 1953.
26. Lindquist, E. F., The Iowa Tests of Educational Development. Science Research Associates, Inc., Chicago, 1952.
27. National Education Association, Educational Policies Commission, Education of the Gifted. Washington, D. C., 1950.
28. Otis, Arthur S., Otis Quick-Scoring Mental Ability Tests, Beta Form. World Book Company, New York, 1954.
29. Prescott, George A., The Development of An Improved Method of Making Capacity-Achievement Comparisons, Unpublished Doctoral Dissertation. Boston University School of Education, Boston, 1950.
30. Sexton, John M., Analyzing Student Achievement in Tenth Grade English, Test Service Bulletin, Number 92. World Book Company, New York, 1959.
31. Science Research Associates, Inc., Iowa Tests of Educational Development Manual for the School Administrator. Chicago, 1957.
32. Sullivan, Elizabeth T., et al., California Short-Form Test of Mental Maturity Manual for Interpreting. California Test Bureau, Los Angeles, 1957.
33. Terman, Lewis M., The Discovery and Encouragement of Exceptional Talent, Test Service Notebook, Number 14. World Book Company, New York, 1954.
34. Witty, Paul, Current Practices in Educating the Gifted Child, Education Monograph. D. C. Heath and Company, Englewood, New Jersey, 1957.

35. Worcester, D. A., The Education of Children of Above-Average Mentality. University of Nebraska Press, Lincoln, Nebraska, 1956.
36. World Book Company, Booklet of Norms, Section III, Modal Age Groups, Metropolitan Achievement Tests. Yonkers-On-Hudson, New York, 1948.