A STUDY OF THE EFFECTS OF GROUP INTERPRETATION OF APTITUDE TEST RESULTS UPON THE ESTIMATES OF ABILITIES AND THE ESTIMATES OF TEST PERFORMANCE OF A GROUP OF SEVENTH GRADE PUPILS

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

C. Lawrence Baymer

1963

This is to certify that the

thesis entitled

A STUDY OF THE EFFECTS OF GROUP INTERPRETATION OF APTITUDE TEST RESULTS UPON THE ESTIMATES OF ABILITIES AND THE ESTIMATES OF TEST PERFORMANCE OF A GROUP OF SEVENTH GRADE PUPILS

presented by

Charles Lawrence Beymer

has been accepted towards fulfillment of the requirements for

Ph.D. degree in Education

James W Costar
Major professor

Date September 10, 1963

O-169

ABSTRACT

A STUDY OF THE EFFECTS OF GROUP INTERPRETATION OF APTITUDE TEST RESULTS UPON THE ESTIMATES OF ABILITIES AND THE ESTIMATES OF TEST PERFORMANCE OF A GROUP OF SEVENTH GRADE PUPILS

By C. Lawrence Beymer

This study was designed in order to develop an approach to the communication and interpretation of test scores and to use the procedure with a group of seventh grade pupils in order to investigate its effectiveness. The problem investigated was the effect of group interpretation of aptitude test scores upon the self-estimates of abilities and the self-estimates of test results of a sample of seventh grade pupils.

The Academic Promise Tests, an aptitude battery measuring abstract reasoning, numerical, verbal, and language usage abilities, were administered to seventh grade pupils in one Indiana junior high school. One month later interpretation sessions were held in regular classroom groups. After a short discussion of individual differences, abilities, and norm groups, subjects estimated their abilities in the areas measured by the tests and their test results. Estimates were made using a chart composed of numbers and 100 cartoon symbols representing 100 boys and girls forming a stanine distribution. Group interpretation of test scores followed immediately, using the test publisher's printed profile. Ten interpretation groups averaged 28.1 pupils each. One week later subjects re-estimated their abilities and test results in the same manner as before.

Data for the study were collected from 244 pupils (119 girls and 125 boys) who took the tests and made both preinterpretation and

postinterpretation estimates. Means and standard deviations were computed for these estimates made by the total group and these sub-groups: boys; girls; high academic aptitude pupils; low academic aptitude pupils; over-estimators of abilities; under-estimators of abilities; over-estimators of test results; and under-estimators of test results. The t-test technique was used to test the significance of difference between means, the F-test to test significance of difference between variances, and chi-square to test the significance of proportions of boys and girls in the various groups.

Major findings are listed below.

1. Postinterpretation mean ability estimates were significantly closer to points of maximum estimation accuracy than were preinterpretation means for these groups: high and low academic aptitude; over-estimators and under-estimators of abilities and under-estimators of test results.

2. Postinterpretation mean test results estimates were significantly closer to points of maximum estimation accuracy than were preinterpretation means for these groups: high and low academic aptitude; over-estimators and under-estimators of abilities; and over-estimators and under-estimators of test results.

3. Highly significant decreases in variances of both abilities and test results estimates were found for the total group and every sub-group in the study after interpretation.

4. For these groups, variances of postinterpretation estimates of test results were significantly less than variances of preinterpretation estimates of abilities; the total study group; boys; girls; high academic aptitude group; under-estimators of abilities; and under-estimators of

test results.

5. No sex differences in accuracy of estimation were found.

6. Initially, high academic aptitude pupils underestimated, low academic aptitude pupils overestimated. After interpretation, both groups moved toward more accurate estimations, the high group close to the points of maximum estimation accuracy.

7. Postinterpretation estimate means of both under-estimators of abilities and under-estimators of test results were significantly closer to the points of maximum estimation accuracy than were postinterpretation estimate means of over-estimators of abilities and over-estimators of test results.

8. The high academic aptitude group had significantly less variance in postinterpretation estimates of abilities and of test results than did the low academic aptitude group.

9. The under-estimators of abilities and the under-estimators of test results had significantly less variance in postinterpretation estimates of test results than did the over-estimators of abilities and the over-estimators of test results.

A STUDY OF THE EFFECTS OF GROUP INTERPRETATION OF APTITUDE TEST RESULTS UPON THE ESTIMATES OF ABILITIES AND THE ESTIMATES OF TEST PERFORMANCE OF A GROUP OF SEVENTH GRADE PUPILS

By

C. Lawrence Beymer

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

College of Education

1963

ACKNOWLEDGMENTS

No doctoral dissertation has ever been the product of one lone individual, and this one is no exception. Many people have contributed much toward its development from genesis to final completion. Thanks and appreciation are due the following: Dr. John Krumboltz of Stanford University, for his criticisms and suggestions at the early design stage; Don Wilkening of the Michigan State University Audio-Visual Department for his assistance with the stanine distribution chart; Principal Nick Martinez and his faculty of the Waverly Junior High School near Lansing, Michigan, for permitting numerous try-outs of the interpretation process; Dr. James Rentschler, Principal of Woodrow Wilson Junior High School in Terre Haute, Indiana, and his faculty, who were not only willing and cooperative, but also eager to have the investigation conducted in their school; Dr. John Patterson of Purdue University, who contributed several suggestions concerning design and statistical treatment of the data; Dr. Charles Hardaway, Director of Research and Testing at Indiana State College, Terre Haute, Indiana, who offered many valuable suggestions during the project; Miss Janet Stephens, who tabulated the data; Mrs. Janet Kessinger, who typed and retyped the text several times, and last but not least, my committee, Dr. James W. Costar, chairman, Dr. Walter F. Johnson, Dr. Gregory A. Miller, and Dr. Carl F. Frost, who allowed me sufficient freedom and flexibility to permit the development of a required task into a valuable learning experience.

-CLB

TABLE OF CONTENTS

CHAPTER PAGE

CHAPTER PAGE

LIST OF TABLES

TABLE PAGE

TABLE PAGE

TABLE PAGE

TABLE PAGE

TABLE PAGE

CHAPTER I

INTRODUCTION

I. PURPOSES OF THE STUDY

The purposes of this study were: (1) to develop an approach to the communication and interpretation of test scores, and (2) to use the procedure with a group of seventh grade pupils in order to investigate its effectiveness.

II. GUIDANCE FOUNDATIONS

Practically non-existent until the turn of this century, guidance and counseling in the schools has become established as one of the basic components of our educational system. Because a wide variety and combination of activities have evolved to meet local needs, attempting to generalize about guidance and counseling practices is difficult. However, even though specific practices and organizational patterns vary from school to school and from state to state, the goals are usually the same. Simply stated, they are: to assist the individual student to understand himself, his environment, and to make wise decisions in light of these understandings. A "wise choice" is defined as one the individual makes for himself after considering the relevant factors and alternatives, a choice for which he is willing to take the consequences. This is the philosophical framework for the present study.

III. THE GROWTH OF STANDARDIZED TESTING

Despite the above-mentioned diversity in pattern and practice psychological testing is undoubtedly one of the most widespread and universal components of guidance programs in our schools. For documentation of this contention, there are figures from many sources which reveal the great number of standardized tests and answer sheets purchased each year. It has been estimated that in 1944 more than 26 million tests were administered by educational institutions, business firms, and personnel consultants to more than 11 million individuals.[1] Another source[2] estimates 1954 sales at 61 million and the number for the 1958-1959 school year at 122 million. In the academic year beginning in September, 1961, and running through June, 1962, the estimated number of standardized tests administered in schools had risen to 125 million, a figure just shy of three times the total number of all students enrolled in public and private education from kindergarten through graduate school.[3] The 1961 Annual Survey of Standardized Test Publishers prepared by Stanley B. Hunt and Associates for the American Textbook Publishers Institute[4] in

---

[1]"Testing is Big Business," American Psychologist, II, January, 1947, p. 26.

[2]"Testing: Can Everyone Be Pigeonholed?" Newsweek, 54, 3, July 20, 1959, pp. 91-93.

[3]"Are Americans Over-Testing?" Overview, August, 1961, pp. 31-33.

[4]1961 Annual Survey of Standardized Test Publishers, prepared by Stanley B. Hunt and Associates for the American Textbook Publishers Institute, 432 Park Avenue South, New York City, New York, April, 1962.

April, 1962, contains the following statistics on net sales of standardized tests and answer sheets.

1954	81,526,000
1955	83,800,000
1956	91,070,000
1957	97,810,000
1958	109,710,000
1959	133,620,000
1960 - Total	140,750,000
Grades 1 - 8	81,800,000
Grades 9 - 12	40,850,000
Grades 13 & over	12,400,000
Industrial	5,700,000
1961 - Total	141,100,000
Grades 1 - 8	81,850,000
Grades 9 - 12	41,950,000
Grades 13 and over	11,550,000
Industrial	5,750,000

Although figures from these sources differ slightly, it seems obvious that testing with standardized instruments has become a popular practice in our schools. Gibson[5] administered questionnaires to 904 high school students in 12 schools of three states. Of this number, only three individuals reported that they had not taken such tests.

Hardaway, Kozak, and Ederle[6] provide the most recent survey data on the extent and cost of testing in the secondary schools of one state,

---

[5]Robert L. Gibson, "Pupil Opinions of High School Guidance Programs," Personnel and Guidance Journal, XL, 5, January, 1962, pp. 453-457.

[6]Charles W. Hardaway, Grace B. Kozak, and Helen Ederle, Budgetary and Financial Provisions for Guidance Services in Indiana Public Secondary Schools, Office of Research and Testing, Indiana State College, Terre Haute, Indiana, 1963.

Indiana. From 661 questionnaires mailed out, 323 usable returns were received, representing 45% of all secondary schools, public and private, in the state. A further breakdown revealed that 10% of the state's small schools (enrollment, less than 100), 46% of the medium-sized schools (enrollment, 100 to 249), and 78% of the large schools (enrollment 250 or more) responded. Data pertinent to this study are as follows:

	small	medium	large	total
Number reporting use of tests (% of schools responding)	73%	83%	91%	86%
Annual per pupil expenditure for standardized tests	93¢	66¢	98¢	95¢

IV. FUNCTIONS OF TESTS IN GUIDANCE AND COUNSELING

Information that standardized tests can supply have many possible uses in the school. Thorndike and Hagen[7] suggest three kinds of functions: classroom, administrative and guidance. Examples of classroom functions include identifying pupils who need special diagnostic study and evaluating discrepancies between potentiality and achievement after determining reasonable achievement levels for each pupil. Typical administrative functions are forming and assigning to classroom groups, and evaluating curricular offerings, emphases, and experiments. Of special interest to this study are the guidance functions of standardized tests which emphasize helping pupils set educational and vocational goals, make

---

[7]Robert L. Thorndike and Elizabeth Hagen, Measurement and Evaluation in Psychology and Education, Second Edition (New York: Wiley & Sons, 1961), p. 446.

immediate choices, and build realistic self-pictures. Thus test results have the potentiality for facilitating progress toward the important goal mentioned earlier: helping individuals make wise decisions through the use of information about themselves and their environments. Such a function rests squarely upon an acknowledgment of the importance of an individual's concept of self as a determinant of his behavior.

V. THE SELF-THEORY FRAME OF REFERENCE

Self-theorists consider an individual's self-concept to be of crucial importance. Combs and Snygg[8] state that what a person thinks and how he behaves are largely determined by his conceptions of himself, including his abilities. Bordin[9] says that the individual responds to choices in terms of his conception of himself, and Super[10] uses the implementation of the self-concept as the framework for his theory of occupational choice and vocational development.

It is generally agreed that this concept of self is not inborn, but must be achieved. Rogers[11] says that the self and personality

---

[8]Arthur Combs and D. Snygg, Individual Behavior, Revised Edition, (New York: Harper & Brothers, 1959).

[9]E. S. Bordin, "A Theory of Vocational Interests as Dynamic Phenomena," Educational and Psychological Measurement, 3, 1943, pp. 49-65.

[10]Donald E. Super, "Vocational Adjustment: Implementing a Self-Concept," Occupations, XXX, November, 1951.

[11]Association for Supervision and Curriculum Development, National Education Association, Perceiving, Behaving, Becoming (Washington, D.C.: N. E. A., 1962).

emerge from experience. Jersild[12] writes that the self-concept develops as the person meets up with the experiences of life. This point of view is supported by White[13] and Combs[14].

Several authorities have taken the position that the establishment of the self-concept is one of the major tasks of the adolescent period. Erikson[15] considers the establishment of identity as the main developmental task of adolescence. Friedenberg[16] says that the central growth process in this age period is to define the self through clarification of experience and to establish self-esteem.

If the self-concept is of crucial importance in determining behavior, if it is learned from experience, and if much of this construction occurs during adolescence, then several implications for educational practice can be identified. Jersild has written that:

> ..... the most important task for child psychology and for education is to find out how the educational program from nursery school through college might help the growing person to understand and accept himself.[17]

---

[12]Arthur T. Jersild, In Search of Self, (New York: Bureau of Publications, Teachers College, Columbia University, 1952).

[13]Robert W. White, The Abnormal Personality (New York: The Ronald Press, 1956).

[14]Arthur W. Combs, "A Perceptual View of the Adequate Personality," Perceiving, Behaving, Becoming (Washington, D. C.: N. E. A., 1962).

[15]E. H. Erikson, Childhood and Society (New York: Norton, 1955).

[16]Edgar Z. Friedenberg, The Vanishing Adolescent (New York: Dell Publishing Company, 1962).

[17]Arthur T. Jersild, In Search of Self (New York: Bureau of Publications, Teachers College, Columbia University, 1952), p. 3.

Peters and Farwell[18] suggest that perhaps the most important discovery that boys and girls need is to make a discovery of themselves. Super supports this view specifically: "Here, then, is a major goal for education: the development of clear, well-formulated, and realistic self-concepts."[19] The guidance services, a part of the total educational program, share this responsibility. Friedenberg states: "The purpose of guidance, after all, is to help students see themselves clearly and realistically, and to accept what they see at least as good enough to go on from."[20]

It appears that several specific kinds of things can be done to help students develop realistic self-concepts.

First of all, we can direct our emphasis toward the assets of each individual instead of toward his liabilities. Combs[21] says that people learn that they are able not from failure, but from success. The writers of the 1962 A. S. C. D. Yearbook[22] point out that often it is the child who feels unwanted and unable who cannot afford to be accurate in self-assessment.

---

[18]Herman J. Peters and Gail F. Farwell, Guidance: A Developmental Approach (Chicago: Rand McNally & Company, 1959).

[19]Donald E. Super, The Psychology of Careers (New York: Harper & Brothers, 1959), p. 111.

[20]Edgar Z. Friedenberg, The Vanishing Adolescent (New York: Dell Publishing Company, 1962), p. 144.

[21]Arthur W. Combs, "A Perceptual View of the Adequate Personality," Perceiving, Behaving, Becoming (Washington, D. C.: N.E.A., 1962), pp. 50-64.

[22]Association for Supervision and Curriculum Development, National Education Association. Perceiving, Behaving, Becoming (Washington, D.C.: N. E. A., 1962), p. 120.

Second, if we really want people to develop unique, personal self-concepts, we must not only accept but encourage variability.

Third, the individual must be accepted by others, including his teachers and counselors, if he is to learn to accept himself. Rogers and his co-workers have demonstrated that being accepted leads to self-acceptance and that, in turn, leads to acceptance of others.

Fourth, we should make available to each student all of the valid and reliable information about himself that he needs in order to make decisions about his own future. This includes the results from standardized tests, the particular focus of this study.

VI. THE UTILIZATION OF TEST RESULTS

While some may believe that test results are for the private use of the expert and for no other eyes, most authorities disagree with this proposition. The authors of the 1962 A. S. C. D. Yearbook take the following position.

> Any information, including achievement and intelligence test ratings, which contributes to the accuracy of his view of self should be available to the student. Withholding records of performance, test results, inventories, or other data which can be made available means withholding information important to decision making.[23]

Tyler[24] states that the most important principle in testing

---

[23]Association for Supervision and Curriculum Development, National Education Association, Perceiving, Behaving, Becoming (Washington, D.C.: N. E. A., 1962), p.132.

[24]Leona E. Tyler, The Work of the Counselor, Second Edition (New York: Appleton-Century-Crofts, 1961), p. 106.

is that information obtained from guidance testing is for the use of the client himself, and that our aim should be to enable the pupil to form a sound idea of his own assets and liabilities. She feels that it is better to do no testing at all rather than to open the door for the many kinds of misconceptions and anxieties that testing without interpretation creates.

Dr. Dorothea McCarthy of Fordham University seems to feel quite strongly about the desirability of interpreting test results:

> For a long time tests were administered in schools and the practice was considered generally desirable, but the information often remained in files on cumulative record cards and was not used or interpreted to the testee himself or to his parents, teachers, or counsellors. Such testing programs are utterly useless and a waste of time and money unless the test results are to be made known and unless they are to be used in decision making about the individuals who have been tested.[25]

Such comments are not new; one of the pioneers of psychological testing, E. L. Thorndike, was saying about the same things in 1924:

> A highly valuable possibility of service for tests, it seems to me, has been but little recognized. Too often tests are given, the data are tabulated, conclusions drawn therefrom are utilized by supervisors, and methods are revised by teachers because of them; but the pupils who write the tests are not informed of any of the results except in those rooms where unsatisfactory conditions have brought about attempts to shift the blame to these pupils.
>
> The final justification for every testing regime rests in Mary Jones and John Smith, and it therefore behooves all persons who are making and giving tests to take them into partnership as soon as and as completely as is feasible.[26]

---

[25]Dorothea McCarthy, "Ethical and Professional Considerations in Reporting Test Information," Proceedings and Summaries, 23d Annual Meeting, New York State Psychological Association, Inc., May 5-7, 1960, pp. 11-12.

[26]E. L. Thorndike, "Tests and Their Uses," Teachers College Record, XXVI, 2, October, 1924, pp. 93-94.

VII. DISUSES AND MISUSES OF STANDARDIZED TESTS

One of the fundamental goals of guidance is to assist the individual student to understand himself, and results from standardized tests can provide information that will aid in the attainment of this goal. In order to fulfil this potential, however, tests must be chosen, administered, scored, and interpreted correctly. Unfortunately, much disturbing evidence is available to support the contention that far too many tests are improperly handled during one or more of these steps.

For example, scoring for most standardized tests seems to be a simple procedure. Yet Durost[27] reports that from 10% to 50% of the answer sheets coming in to his I. B. M. test scoring service have to be cleaned up with erasers before accurate scoring is possible, and that 25% to 33% of teacher-scored tests coming in have to be completely rescored to provide enough accuracy for mass statistical treatment. Phillips and Weathers[28] tabulated the errors made by teachers in 5,017 scorings of parts of the Stanford Achievement Test, and found that 28% of the tests contained one or more errors in scoring. Test scoring is said to be the easiest and most objective phase in the utilization of the instruments; if frequent and serious errors are discovered at this stage, it is difficult to make

---

[27]W. N. Durost, "Present Progress and Needed Improvements in School Evaluation Programs," Educational and Psychological Measurements, 14, pp. 247-254.

[28]Beeman N. Phillips and Garrett Weathers, "Analysis of Errors Made in Scoring Standardized Tests," Educational and Psychological Measurement, 18, 1958, pp. 563-567.

optimistic inferences about the more complex processes involved.

We have no reason to believe that every student who takes a standardized test later receives an accurate interpretation of the results; in fact, we cannot be sure that even a majority of them receive the results in any form. Hastings and associates[29] surveyed 1,000 eleventh-grade students in 38 randomly-selected high schools in Illinois, obtaining the following responses in percents to the question, "Does the counselor discuss your test results with you?"

Curriculum	Never	Rarely	Sometimes	Frequently
College prep	15	14	42	29
Business-commercial	23	21	40	16
Vocational	26	21	38	15
Other	37	18	45	9

In the previously-mentioned survey by Gibson,[30] his respondents revealed interesting feelings toward test interpretation. When asked if they understood what their test results mean, 43% answered "not sure" and "no." Slightly more than 50% said they would have liked further interpretations of their scores. As far as self-understanding is concerned, only 44% indicated that they understood themselves to their own satisfaction, and approximately the same percentage said they felt their teachers seemed to understand them.

---

[29] Hastings, et al. "The Use of Test Results - Cooperative Research Project Number 509" Bureau of Educational Research, University of Illinois, 1960.

[30] Robert L. Gibson, "Pupil Opinions of High School Guidance Programs," Personnel and Guidance Journal, XL, 5, January, 1962, pp. 453-457.

No interpretation may be preferable to improper interpretation. Barry and Wolf state: "Perhaps the most damaging abuse of testing is the incredibly poor test interpretation done in many schools."[31] They also take the position that nothing is more dangerous to the individual than the mis-interpretation of test results.

Arnold[32] developed a questionnaire about the Kuder Preference Record and administered it to 200 entering college freshmen. He found that 44% of them had taken the Kuder Preference Record while in high school. Examination of the responses led Arnold to conclude that the written statements of what these students said they learned from Kuder Preference Record showed confusion of interests and abilities, and of measured interests with other interests. He discovered that a "considerable number" of these inventories had been given without the results being discussed later with the students. Arnold speculates that inadequate test interpretation in both quantity and quality must take the blame for this situation.

Berg[33] tabulated the misunderstandings he noted in 30 clients of six counselor trainees at Northwestern University. All of these counselors held the M.A. in either psychology or education (guidance), and all of them had some full-time counseling experience prior to entering

---

[31]Ruth Barry and Beverly Wolf, Epitaph for Vocational Guidance (New York: Bureau of Publications, Teachers College, Columbia University, 1962), p. 52.

[32]Dwight L. Arnold, "Student Reaction to the Kuder," Personnel and Guidance Journal, September, 1958, pp. 40-42.

[33]Irwin A. Berg, "Test Score Interpretation and Client Confusions," Personnel and Guidance Journal, May 1956, pp. 576-578.

the practicum course. Types of test score misunderstandings revealed by the clients and their frequencies of occurrence were as follows:

Centile confusion with IQ	13.4%
Confusion over what IQ means	16.7%
Norm group confusion	40.0%
Confusion of interest and aptitude	50.0%
Scores a guarantee of success	56.7%

Berg feels that most of this confusion is the fault of erroneous counselor conceptions and careless interpretations.

Finally, Leo Goldman in his book Using Tests in Counseling takes the following position.

> Finally, there is an alarming amount of misuse and disuse of test results....From the vantage point of this observer, the impression has been that many, if not most, of the tests currently in use are either wasted or, even worse, used in such a way as to misinform and mislead.[34]

VIII. STATEMENT OF THE PROBLEM

With the above factors in mind, this study was designed in order to develop an approach to the communication and interpretation of test scores, and to use the procedure with a group of seventh grade pupils in order to investigate its effectiveness. The problem to be investigated in this study is the effect of group interpretation of aptitude test scores upon the self-estimates of abilities and the self-estimates of test results of a sample of seventh grade pupils.

IX. SUMMARY

The extent of standardized testing has more than kept pace with

---

[34]Leo Goldman, Using Tests in Counseling, (New York: Appleton-Century-Crofts, 1961), p. 2.

the growth of guidance and counseling practices in our schools. Potentially, results from such instruments can provide students with information valuable in the formation of accurate self-concepts, but available evidence suggests that most of the results of such tests now being administered are disused if not actually misused. The purpose of this study was to develop a test interpretation procedure which would increase the effectiveness of the interpretation process.

The following chapter presents a comprehensive review of the related literature. Chapter III contains an account of the design and methodology of the study, followed by a presentation of the findings in Chapter IV. A summary of the study, the conclusions of the study, and some recommendations for both practice and further research appear in Chapter V.

CHAPTER II

A REVIEW OF RELATED LITERATURE

I. RECENT RESEARCH

Until fairly recent times, directions for test interpretation have explicitly or implicitly stated that it is to be done in the one-to-one relationship of the counseling interview. Some writers in the field still feel that this is the only acceptable approach. Barry and Wolf, for example, typify this point of view:

> Group interpretation of tests is usually even more disasterous than the ordinary individual approach. The teacher or counselor attempting the interpretation cannot take into account all the feelings and values of thirty-odd students. In an attempt to handle an impossible situation, he is likely to become either so vague as to be unintelligible or so authoritarian as to eliminate discussion. . . . Students are not, and should not be expected to be experts in measurement. It is the counselor's responsibility to help the students to develop some comprehension of what their own scores mean.[1]

Nevertheless, in the past few years interpretation of test results to groups of students seems to have become more and more common. Perhaps the original reason was an attempt to utilize valuable and limited counselor time more efficiently; any necessary compromise in effectiveness was thought to be a better alternative than no interpretation at all. But some workers in the field, probably beginning with Froehlich, began to suggest that not only was group interpretation just as effective as individual interpretation in many cases, but perhaps even more so. A search through the literature of the past

---

[1]Ruth Barry and Beverly Wolf, Epitaph for Vocational Guidance, (New York: Bureau of Publications, Teachers College, Columbia University, 1962), pp. 53-54.

decade does not produce a great number of studies of group interpretation, but it must be added that a similar search would not disclose many studies dealing with any type of test interpretation. The following seem to be the major ones that have produced implications for this particular study.

Froehlich and Mosher[2] studied what they considered to be a logically expected outcome of counseling, the memory of test scores. They did group and individual interpretations of Differential Aptitude Tests scores with 150 ninth graders, which include having them draw their own profiles which they kept. Fifteen months later, the experimenters found that a large proportion of the students did not report accurately when they were asked to redraw their D. A. T. profiles. Although high scores were remembered more accurately than low ones, there was a tendency for both high and low scores to be reported as being closer to the mean than they actually had been. The authors contend that this finding contradicts the "general belief" among counselors that it is easier for high rather than low ranking pupils to understand and accept test evidence, if one assumes that accuracy of self-report fifteen months later is closely related to understanding and acceptance.

Two evaluative comments seem applicable. First, the unequal size characteristic of percentiles may have functioned to make memory

---

[2]C. P. Froehlich and W. E. Moser, "Do Counselees Remember Test Scores?" Journal of Counseling Psychology, I, Fall, 1954, pp. 149-152.

appear worse than it actually was. For example, is it fair to say that the same degree of memory inaccuracy is demonstrated by the student who remembers his score as the 80th percentile when it acutally was the 90th, and the student who remembers his score as the 50th percentile when it actually was the 60th? Secondly, the fifteen-month time interval actually represented a period starting early in the freshman year in high school to about a third of the way through the sophomore year. During this period of time the students might have grown in maturity, competency, achievement, and self-understanding to the point that the profile drawn then was the one they felt applied to them at that time, and not a reproduction of something that might have been true of them over a year previously.

Lane[3] did individual interpretations with 111 high school students, comparing the effectiveness of what he called the traditional, directive, counselor-centered method with that of a permissive, non-directive, client-centered method. He found no significant differences. Two criterion measures of self-understanding were used, a check list and an essay, but the experimenter reported disappointment at the low level of performance on the essays. Low positive relations were found between check list scores and measures of scholastic aptitude, socio-economic background, and even less for social adjustment, although all were declared "minor influences."

---

[3]David Lane, "A Comparison of Two Techniques of Interpreting Test Results to Clients in Vocational Counseling," (unpublished doctoral dissertation, Columbia University, 1952).

Lallas[4] compared three methods of interpreting achievement test scores: (1) individual counseling interviews, (2) group interpretation, and (3) a combination of the individual plus group approaches. His high school junior subjects estimated their rankings on the various subtests of the Iowa Tests of Educational Development before and after interpretation of their own scores. All three experimental groups showed significantly greater improvement in accuracy of self-estimate than the control group, with greatest improvements appearing in the individual counseling and the group-plus-individual counseling groups, with somewhat less improvement shown by those receiving only group interpretations.

Generalization and application of the results of this experiment are limited by weaknesses in the design and procedure. The experimenter attempted to match subjects in his experimental groups; the group-plus-individual counseling experimental group received much more attention than the others, and seven different graduate student counselors did the individual interviews while one individual did all of the group interpretations. Specific details of what actually transpired in the group interpretation sessions is missing, and only vague details of the individual interview procedure is given.

Wright,[5] working with college students, found that while both individual and multiple or group interpretations resulted in significantly

---

[4]John E. Lallas, "A Comparison of Three Methods of Interpretation of Achievement Tests to Pupils," (unpublished doctoral dissertation, Sanford University, 1956).

[5]Edward Wayne Wright, "A Comparison of Individual and Multiple Counseling in the Dissemination and Interpretation of Test Data," (unpublished doctoral dissertation, University of California, 1957).

more accurate self-reports than no interpretation at all, he could demonstrate no meaningful differences between the individual and group approaches. Unfortunately, details of the procedures used in either method are not given.

Froehlich,[6] in one of the rare published studies that utilized high school students, concluded that his data did not support the claim that counseling (which in this case involved the interpretation of the Differential Aptitude Tests) must be individual. Using several methods of data analysis, he found no significant differences between individual and group interpretation in all methods but one, and in that instance the group approach was significantly more effective. His design seemed to have several weaknesses, however, including non-randomization of volunteer subjects, varying group size, and only a total of 42 subjects involved. No details or descriptions of either the individual or group interpretations are available.

Gustad and Tuma,[7] using both client-centered and more directive techniques of test score interpretation with male undergraduate college students, found no significant differences in client learning about self. Moreover, they could demonstrate no significant relationship between scholastic aptitude and client learning about self during counseling.

---

[6]Clifford P. Froehlich, "Must Counseling Be Individual?" Educational and Psychological Measurement, XVIII, 4, 1958, pp. 681-689.

[7]John W. Gustad and Abdul H. Tuma, "The Effects of Different Methods of Test Introduction and Interpretation on Client Learning in Counseling," Journal of Counseling Psychology, Vol. 4, No. 4, 1957, pp. 313-317.

They did find that initial accuracy of self-ratings was positively related to client learning, suggesting that those who knew themselves best before counseling learned the most about themselves during counseling, assuming that the test scores involved represent the "truth" about their interests and abilities.

L. B. Rogers[8] found that a self-evaluative technique in which non-test data were surveyed before test results were introduced and in which client participation was encouraged was no more effective than a test centered method in which the profile was explained in detail but no particular effort was made to stimulate client participation or to introduce non-test data. His college student subjects given the one type of interview showed no greater growth in self-understanding than those given the other kind, when both groups included a wide range of intelligence and of responsiveness in the interview.

Some modest support for the value of greater client participation was found when each of the two main treatment groups was subdivided into two groups, those who had participated actively in the interview and those who had not. The sub-group which participated actively in the self evaluative interview showed a significant increase in self-knowledge; others increased insignificantly.

The factor of intelligence as measured by the A. C. E. Psychological Examination was also studied. When the highest and lowest four deciles in each of the two main treatment groups were studied, it was found that

---

[8]Lyle B. Rogers, "A Comparison of Two Kinds of Test Interpretation Interview," Journal of Counseling Psychology, 1954, 1, pp. 224-231.

the self-evaluative approach was as effective with the less intelligent as with the more intelligent, whereas the test-centered approach was not as effective with the less intelligent as it was with the more intelligent.

Evidence concerning the role of the counselee in test interpretation is provided by a study conducted by Dressel and Matteson at Michigan State.[9] Forty recorded interviews with freshmen college students with seven counselors were used. They tested the hypotheses that in comparison with counselees who participate less in the test interpretation process, those who participate more actively (1) gain more in self-understanding, (2) are more certain of their vocational goals, and (3) are more satisfied with the experience. Their findings were moderately supportive of the first and second hypotheses, but not of the third.

Greater gains in self-understanding were made by counselees whose counselors succeeded in eliciting the greatest amount of counselee participation in the interview, and they were more certain of their vocational goals afterward. However, their results do not indicate that clients who participate more get more out of the experience regardless of who the counselor is and what he does.

Of particular importance to this study are the findings of an investigation conducted by Kamm and Wrenn at Minnesota.[10] They set out

---

[9]Paul L. Dressel and Ross W. Matteson, "The Effect of Client Participation in Test Interpretation," Educational and Psychological Measurement, X, 4, Winter, 1950, pp. 693-706.

[10]Robert B. Kamm and Gilbert C. Wrenn, "Client Acceptance of Self-Information in Counseling," Educational and Psychological Measurement, X, 1, Spring, 1950, pp. 32-42.

to study within which interview situations counselees do and do not tend to accept presented information, how those who do accept information differ from those who do not, and what type of information does and does not tend to be accepted. "Acceptance" was defined as favorable reception by the client of information presented to him, as demonstrated by what the counselee said and did. "Information" was defined as all data presented by the counselor, whether they be in the form of advice, suggestions, emphases, recommendations, interpretations, requests, or explanations. Forty recorded educational-vocational planning interviews with one trained, experienced counselor were analyzed, with additional data coming from post-interview check lists and follow-up interviews one and four months later.

They found that both acceptance and non-acceptance of information occured in situations in which the counselee-counselor relationship was judged to be friendly, suggesting that evaluation of effectiveness cannot be done validly by measuring good will.

With the exception of information involving alternation of previously-made counselee plans which tended to be more often accepted by the acceptance group, different kinds of information were accepted equally well by acceptance and non-acceptance groups.

No relationship was found between acceptance of information and any of the following factors: academic aptitude, particular measured personality patterns, social class, veteral status, marital status, previous counseling interviews, length of interview, or proportion of time which the client spoke during the interview.

They found that information is likely to be accepted if it is presented in an emotionally-relaxed atmosphere; is directly related to the counselee's own immediate problem and felt needs; and is congruent with and not in opposition to his self-concept. The authors state that perhaps the most conclusive of all their findings is that the client himself is the basic determiner of whether or not acceptance occurs. They suggest that the feelings, needs, wants, desires, and attitudes of the counselee are more important than the characteristics of the interview situation or type of information presented.

The authors recommend that counselors use techniques designed to assist in the development of the counselee toward a more realistic awareness of himself, and that they be aware of the level of thinking of their counselees.

Singer and Stefflre[11] investigated the effects of counseling upon the expressed self-interests of high school seniors, analyzing both means and standard deviations for significance. For the boys, they found no significant change in means, and for the girls, significant mean changes only for the Science and Mechanical scales. When standard deviations were compared, it was found that all of the standard deviations for the boys decreased, although only for the Mechanical scale was this decrease significant. For the girls, four of the six standard deviations decreased, with significant decreases on the Science and Mechanical scales. Thus

---

[11]Stanley L. Singer and Buford Stefflre. "Analysis of the Self-Estimate in the Evaluation of Counseling," Journal of Counseling Psychology, I, No. 4, Winter, 1954, pp. 252-255.

only five of a total possible twenty-four comparisons were statistically significant.

II. LIMITATIONS OF PREVIOUS STUDIES

The following are limitations of the studies of test interpretation which have been reported:

1. In most cases only a small number of subjects was used.

2. In most cases sampling methods violated the most fundamental rules. More than one utilized volunteers, and randomization of subjects to various treatments was seldom mentioned.

3. Group size varied greatly, sometimes within the same study. Seldom was group size near what is considered to be a typical classroom group. Few groups in available studies approach as many as a dozen subjects at a time.

4. An examination of the original sources produces few details of what was said and done in the interpretation sessions. Specific replications would be most difficult if not impossible.

5. Most of the studies were performed at the college level, only a few used high school students, and none used students at the junior high school level. If concept of self is involved in the communication of test results, it seems that investigations should be undertaken at that stage where the self-concept is being formed.

III. SUMMARY

To summarize, available research pertinent to this problem seems to justify the following generalizations:

1. No consistent superiority of any one particular approach (ex., directive vs. non-directive) has been demonstrated.

2. Techniques of group interpretation have not been clearly demonstrated to be inferior to techniques using individual interpretation.

3. In every available case, individuals who have had their test scores interpreted later demonstrate a higher degree of accurate self-knowledge than those who received no interpretation.

4. Acceptance of test results has not been shown to be dependent upon level of mental ability as measured by traditional standardized academic aptitude tests, measured personality patterns, social class, level of performance, or counselee experience in the counseling setting.

5. Acceptance of test results appears to be related to the status of the counselee's self-concept as it is either threatened or strengthened by the information being presented.

6. Counselors who succeed in eliciting counselee participation in the interpretation process are likely to make possible greater gains in self-understanding by the counselees.

7. Although moderate successes have been demonstrated in communicating the results of testing, a distressingly large proportion of counselees seem to come away from interpretation sessions little if any wiser about their characteristics than before.

CHAPTER III

DESIGN AND METHODOLOGY OF THE STUDY

I. THE PROBLEM

The results of research studies reported in the previous chapter suggest that techniques of group test interpretation can be improved, and also suggest the directions in which further investigations should proceed.

This study was a combination of the experimental and descriptive approaches. Essentially, the purpose of the study was twofold: (1) the development of an approach to the communication and interpretation of the concepts of individual differences, norm groups, and the meanings of test results, and (2) the use of this approach with a large group of seventh grade students in order to gage its effectiveness.

Following procedures reported in previous studies in this area, interpretations were made to groups of subjects who estimated their test results prior to and after this interpretation. The interpreter attempted to remain neutral toward the test data and permissive toward the students' reactions to these data. Pupils were invited to request individual counseling interviews in case they wanted to discuss their results either further or privately.

This study goes beyond previous ones in several ways. It involved students in the seventh grade, an earlier grade level than is reported in the literature. A larger number of subjects was involved. In addition to estimating test results, subjects also estimated their abilities

in those areas sampled by the test. Estimates were based upon a stanine scale, which utilized a combination of numbers and cartoon symbols. (See Appendix B, p.114.) It was felt that this device had many advantages over the traditional thermometer-like percentile graph in communicating the concepts of individual differences and of norm groups. The one hundred cartoon symbols were used in an attempt to visually sumbolize these concepts. Estimates in the form of stanine bands rather than in percentiles were used in order to lessen the temptation to think of a score as a point rather than as a general area, and to eliminate the disadvantage of the unequal units of the percentile distribution. Moreover, stanines can be processed mathematically without having to be transformed statistically.

As another innovation, the presentation of test results was preceded by a presentation and discussion of the concepts of individual differences and of norm or reference groups. It was reasoned that if the pupil had some prior understanding of the natural and expected variability in people, the theory of the normal curve of distribution, and the nature of norm groups, he would be able to accept the test results which might otherwise be rejected.

This experimental approach was designed to involve the individual pupil at every step. The pupils estimated on a mimeographed form their approximate positions in a normal stanine distribution of their peers. (See Appendix C, p. 116.) Test results introduced as the test's estimate of these same approximate positions, may or may not have agreed with the pupil's estimates. In either case, he was free to either alter his thinking about himself or to reject the test results. No value judgment was placed

upon the degree of agreement with test results. That is, the individual who disagreed with the implications of the test results and who indicated this was not judged to be unrealistic in his opinions. Obviously, the test results may have been, for him, somewhat inaccurate. In other words, the assumption was made that test scores can be wrong.

The experimental estimation procedure and interpretation procedure were gleaned from recommendations of authorities in the fields of guidance and testing, from those descriptions that can be found in the original sources of previous studies, and from the implicit and explicit recommendations of the test publisher. It was felt by the experimenter that these methods are consistent with good practice as it is known and accepted at this time. Complete details of what was said and what occurred during the two meetings with each experimental group can be found in Appendix A, page 104, and Appendix D, page 118.

II. LIMITATIONS OF THE STUDY

Several limitations are apparent in this study that have an effect upon the conclusions which can be drawn from the data. The following are recognized shortcomings:

1. Only one week intervened between the date of interpretation and the follow-up collection of data. This period of time was chosen partly by the realities of the situation. However, it is difficult to accurately measure the effect of time on the ability of the individual

subjects to recall information obtained in test interpretation interviews.

2. The experimenter did all of the estimation and interpretation presentations. Therefore the results may have been at least in part a reaction to his personality and techniques.

3. The sample consisted of seventh graders only, selected from a single school. The extent to which the findings of the study can be generalized are therefore limited to the extent to which this sample represents seventh graders in general, and to the extent that the performance of seventh graders is like that of pupils at other levels.

4. Rapport between the investigator and the various groups was an unmeasured factor. While the impression of the experimenter was that rapport was successfully established between himself and the various groups, no objective evidence is available to either support or refute this opinion.

III. SAMPLE SELECTION

Two major factors influenced the decision to conduct this study at the seventh grade level. First, the seventh grade typically coincides with the beginning of adolescence, a period in which a concept of self is said to be emerging as one of the products of experience. The individual's conceptions of his abilities are an integral part of this self concept. Second, as mentioned previously in Chapter II, most of the published

investigations of test interpretation have been conducted with college students, a few with high school pupils, but none at the junior high school level. For these reasons it seemed appropriate to use seventh grade boys and girls as subjects for this study.

All subjects were drawn from the seventh grade class of one junior high school, Woodrow Wilson Junior High School, Terre Haute, Indiana. This school offered a large number of subjects drawn from at least three readily identifiable socio-economic areas of the city. The principal and the faculty were receptive to assisting with carrying out this investigation. Although some guidance services were being provided, because of a lack of counselor time, test interpretation to the pupils was not at that time being done. All students in the seventh grade class were tested. Only those students present on the date of the test, the day of the interpretation, and the day of the follow-up were included in the study. Table I on page 31 summarizes the number of subjects involved in the study. Ten groups of seventh grade pupils averaging 28.1 pupils per group received the interpretations. Complete data were collected from 244 subjects, 119 girls and 125 boys.

TABLE I

DISTRIBUTION OF SUBJECTS INVOLVED IN THE STUDY

Group Number	Number in interpretation group			Number of usable subjects--A.P.T., pre-interpretation est., post interp. est.		
	Boys	Girls	Total	Boys	Girls	Total
7A1	9	16	25	8	15	23
7A2	17	16	33	15	16	31
7A3	15	14	29	12	10	22
7A4	14	13	27	10	12	22
7A5	10	7	17	7	7	14
7A6	16	12	28	14	11	25
7A7	17	13	30	12	11	23
7A8	17	10	27	16	8	24
7A9	13	18	31	12	17	29
7A10	19	15	34	19	12	31
Total	144	137	281	125	119	244

IV. SELECTION OF THE TEST

The test chosen for administration and interpretation was the Academic Promise Tests, an aptitude battery recently published by the Psychological Corporation of New York. In part an outgrowth of the Differential Aptitude Tests, the A.P.T. battery was designed for use in grades six through nine. A brief description of these tests follows; a more complete description appears in the A.P.T. Manual.[1]

Four types of tests compose the battery, each selected with certain principles and objectives in mind. The verbal section was designed to measure the understanding of word meanings and the ability to use words in reasoning. Items are of the analogies type. The numerical section was designed to measure the capacity to think in quantitative terms, to understand and use numerical relationships. Items are of several types, emphasizing the understanding of concepts and reasoning. The reading requirement for this section is low as few words are used. The abstract reasoning section was designed to measure the ability to see relationships and to recognize concepts presented in the form of diagrams or symbols rather than in words or numbers. Items are of the figure classification type, requiring the pupil to seek out the principle which provides a common characteristic for a set of three figures, and to recognize which of several other figures shares that characteristic. The language usage section was designed to measure

---

[1]George K. Bennett, Marjorie K. Bennett, Dorothy M. Clendenen, Jerome E. Doppelt, James H. Ricks, Jr., Harold G. Seashore, and Alexander G. Wesman, Academic Promise Test Manual, 1962, New York: The Psychological Corporation, 1962.

understanding of correct writing and speech, together with the appreciation of good English. Items require the pupil to identify errors in grammar, spelling, and punctuation. Items on all sections are arranged in order of increasing difficulty. Total testing time is ninety minutes.

A common practice in measurement is to group abilities into verbal and non-verbal classifications. With the A.P.T., the verbal and language usage tests can be combined to provide a measure of broad competence with verbal materials. In a similar fashion the numerical and abstract reasoning sections together provide a measure of broad competence with non-verbal materials. Finally, all four scores when added together provide a single statement of general academic aptitude.

Norms for the A.P.T. are based upon the performances of more than 34,141 pupils in grades 7-9, including 9,141 seventh graders. This norms group was put together by a careful stratified sampling procedure which took into account pupils enrolled by grade level, geographic region of the nation, and community size.

A summary of validity data is given in Table II on page 34 and a summary of reliability data appears in Table III on page 35.

TABLE II

MEDIAN VALIDITY COEFFICIENTS BETWEEN A.P.T. SCORES AND SCHOOL GRADES[2]

	AR	N	V	LU	AR+N	V+LU	A.P.T. Tot.
English	.31	.50	.45	.60	.45	.59	.58
Mathematics	.30	.58	.36	.41	.54	.46	.56
Social Studies	.29	.45	.41	.45	.40	.51	.54
Science	.33	.49	.46	.46	.48	.50	.56

[2]George K. Bennett, Marjorie K. Bennett, Dorothy M. Clendenen, Jerome E. Doppelt, James H. Ricks, Jr., Harold G. Seashore, and Alexander G. Wesman, Academic Promise Test Manual, 1962, Tables 4, 5, 6, and 7. New York: The Psychological Corporation, 1962.

TABLE III

RELIABILITY DATA FOR THE ACADEMIC PROMISE TESTS[3]
GRADE 7

	AR	N	V	LU	AR+N	V+LU	A.P.T. Total
Alternative-form reliability coefficients	.82	.87	.82	.88	.88	.90	.93
Standard error of measurement in points of raw score (maximum score on each of the first four sections = 60)	5.1	3.4	3.6	3.7	6.7	5.6	8.7

[3]George K. Bennett, Marjorie K. Bennett, Dorothy M. Clendenen, Jerome E. Doppelt, James H. Ricks, Jr., Harold G. Seashore, and Alexander G. Wesman, Academic Promise Tests Manual, 1962, Tables 23 and 24. New York: The Psychological Corporation, 1962.

It was decided to report to the pupils only the verbal, numerical, language usage, and abstract reasoning scores, and to ignore for the purposes of this study any consideration of either the verbal, non-verbal, or total A.P.T. scores. The reasons for this decision were as follows:

1. While the four basic abilities measured are relatively concrete, objective, and easy to explain and understand, the combinations of them are more abstract, subjective, and difficult to explain and understand.

2. The total A.P.T. score can be interpreted in much the same manner as the score from "mental ability" or "group intelligence" tests. It was felt that the introduction of such concepts might introduce unwanted and contaminating emotional factors into the learning situation.

3. It seemed that four factors were enough to deal with in an introductory investigation of test score interpretation at this grade level.

V. COLLECTION OF THE DATA

On Monday, February 4, 1963, all seventh grade pupils of Woodrow Wilson Junior High School, Terre Haute, Indiana, took Form A of the Academic Promise Tests. These tests were administered in the pupils' first period classrooms by their regular first period teacher and a college student assistant trained by the experimenter in group testing procedures. During the next month, answer sheets were scored, individual profiles were prepared, and group statistical analysis was begun.

On Monday, March 4, 1963, and Tuesday, March 5, 1963, the experimenter returned to the school for the second phase of the study. This step

required one class period per group. A complete record of what was said and done appears in Appendix A, pages 104 to 112. Using a chart illustrating a norm group (Appendix B, page 114), and an estimation worksheet (Appendix C, page 116), each pupil made an estimate of his or her abstract reasoning ability, numerical ability, verbal ability, and language usage ability. Also each pupil estimated his or her test score in these areas. The two sheets were then collected. The final step of this second phase of the investigation was to interpret the test results, using the test publisher's profile sheets which had been drawn previously for each pupil. (A complete record of this interpretation presentation appears in Appendix A, pages 104 to 112).

Exactly one week later, the experimenter returned to each group for the third phase of the investigation. Following the procedure presented in Appendix D, pages 118 to 123, and using the norm group chart (Appendix B, page 114), and another estimation worksheet (Appendix C, page 116), the pupils re-estimated their test scores and abilities in the four areas sampled by the test.

Thus each student contributed the following data: before he received interpretations of his test scores, stanine estimates of his abstract reasoning ability, numerical ability, verbal ability, and language usage ability, together with an estimate of his test scores in each of these areas, plus the same estimates one week after his test scores had been interpreted. Since estimates were on a stanine scale, four numbers made up each of the two preinterpretation estimates, and four composed each of the two postinterpretation estimates.

VI. ANALYSIS OF THE DATA

Computation of the estimation scores. Raw scores for each of the sub-tests of the A.P.T. were converted first to percentiles (based upon the national norms of the test publisher), and then to stanines. The discrepancy of each estimate stanine from the actual test score stanine was computed, and the four numbers for each of the four estimates were summed algebraically. The following hypothetical data will serve as an example.

		Preinterpretation Estimates				Postinterpretation Estimates			
		AR	N	LU	V	AR	N	LU	V
Abilities	Test Scores	9	7	5	4	9	7	5	4
	Estimates	5	5	5	5	8	6	5	5
		-4	-2	0	+1	-1	-1	0	+1
			= -5				= -1		
Test Results	Test Scores	9	7	5	4	9	7	5	4
	Estimates	6	6	5	5	9	8	5	4
		-3	-1	0	+1	0	+1	0	0
			= -3				= +1		

In order to avoid minus numbers, a constant of +32 was added. This sum was selected because the maximum under-estimation possible is -32. Such an event could occur if an individual had four stanine test scores of 9, but made four estimates of 1, thus scoring 4(-8) on the scale. Likewise, the maximum over-estimation would be 32, which could occur if

an individual had four stanine test scores of 1, but made four estimates of 9, thus scoring 4(+8) on the scale. With the constant added, the final scale runs from 0 to 64, with 32 marking the point of no discrepancy of estimates from test scores. Scores above 32 indicate over-estimations, scores below 32, under-estimations.

In the example above, adding the constant of +32 to each of the four preliminary estimation sums produces the following estimation scores: preinterpretation estimate of abilities: 27; preinterpretation estimate of test results: 29; postinterpretation estimate of abilities: 31; postinterpretation estimate of test results: 33. All data for an individual were in this manner reduced to four numbers, one for his pre-interpretation estimate of his abilities, one for his preinterpretation estimate of his test scores, one for his postinterpretation estimate of his abilities, and one for his postinterpretation estimate of his test scores. These are the data that were analyzed statistically for significance of change.

An arbitrary decision was made by the experimenter not to investigate in this study estimation changes for the various individual abilities and individual tests of the A.P.T., or to compare the accuracy of estimation of various combinations of these individual abilities and individual tests. That is, it was decided not to pursue such questions as how the pupils performed in estimating their verbal ability or if there was any meaningful difference in their accuracy of estimating verbal test scores as compared to abstract reasoning test scores. The rationale for this decision is as follows.

First, it should be noted that the four abilities under discussion here are not completely separate and independent mental factors. Neither are the various tests of the A.P.T. completely separate and independent. Table IV on page 41 presents the mean intercorrelation coefficients of the A.P.T. Examination of these data reveals considerable relationship between the individual tests and combinations of tests.

A second factor somewhat related to the first is the relationship of the length of a measuring instrument or procedure to its reliability. It seemed reasonable to assume that a total estimation score formed by summing four abilities which are to some degree related would be more reliable than an estimate of one fourth that size, the estimate of a single ability.

Finally, it was felt that these topics were not the main focus of an investigation which was to be one of the first studies of test score interpretation at this grade level.

TABLE IV

MEAN INTERCORRELATION COEFFICIENTS OF THE A.P.T.[4]

	N	V	LU	AR+N	V+LU
Abstract Reasoning	.55	.51	.37	--	.49
Numerical		.69	.59	--	.70
Verbal			.63	.67	--
Language Usage				.54	--
AR + N				--	.66

[4]George K. Bennett, Marjorie G. Bennett, Dorothy M. Clendenen, Jerome E. Doppelt, James H. Ricks, Jr., Harold G. Seashore, and Alexander G. Wesman, Academic Promise Tests Manual, 1962. Table 26. New York: The Psychological Corporation, 1962.

Data analysis procedures. Means, variances, and standard deviations of the preinterpretation and postinterpretation estimates of abilities and of the preinterpretation and postinterpretation estimates of test results were computed for the following groups: the total study group; boys; girls; the overestimators of abilities; the underestimators of abilities; the overestimators of test results; and the underestimators of test results.

Statistical techniques. Two statistical techniques were utilized in this study, the t-test of significance of difference between means and the F-test of homogeneity of variances.

This investigation focused on the effects of the interpretation of aptitude test results upon the estimates of abilities and the estimates of test results of seventh grade pupils. By means of the t-test, means of the preinterpretation estimates of abilities of the total group and of the various subgroups were compared with the postinterpretation estimates of abilities of the same groups. In a similar manner, the means of the preinterpretation estimates of test results of the total group and of the various subgroups were compared with the postinterpretation estimates of test results of the same groups. This was done in order to assess the probability of significant change in mean estimates.

The F-distribution tests the hypothesis that two normal distributions have equal variances. It is an appropriate technique in this situation because an examination of means only may fail to identify important differences in variability. In this study, for example, one individual or two different individuals might earn estimation scores of 32 in quite

different ways. One such score could be earned by estimating all four test scores accurately. However, it is also possible to earn an estimation score of 32 by balancing two overestimations with two underestimations. Since estimations can have the similar means but quite dissimilar variances, it is necessary to investigate possible significant changes in variances of estimations as well as possible significant changes in estimation means.

In this instance the F-test was used to ascertain the probability of as great a difference (irrespective of direction, i.e., a hypothesis or decision requiring a two-tailed test) between the variances for two groups. McNemar[5] says that where the significance of variability of two groups is being tested, an F at the .01 point of the table means significance at the .02 level; an F at the .05 level means significance at the .10 level; and an F at the .001 level indicates significance at the .002 level.

Level of significance selected. For this particular study it was felt that establishing a minimum level of significance less than .05 was not warranted. The .05 level of significance, therefore, was selected as the required acceptable level of significance for this study. Any difference so large that it would be expected to occur by chance alone only 5 or less times in 100 will be accepted as evidence sufficient to reject the hypothesis of no difference.

---

[5]Quinn McNemar, Psychological Statistics, (New York: John Wiley & Sons, Inc., 1955), p. 246.

VI. HYPOTHESES TO BE TESTED

1. Postinterpretation estimates of abilities will not differ in accuracy from preinterpretation estimates of abilities.

2. Postinterpretation estimates of test results will not differ in accuracy from preinterpretation estimates of test results.

3. Preinterpretation estimates of abilities will not differ in accuracy from preinterpretation estimates of test results.

4. Postinterpretation estimates of abilities will not differ in accuracy from postinterpretation estimates of test results.

5. No difference in accuracy of estimates of abilities or of estimates of test results will be found between the sexes, between the high and low academic aptitude groups, between the over-estimators of abilities and under-estimators of abilities groups, or between the over-estimators of test results and the under-estimators of test results group.

VIII. SUMMARY

The Academic Promise Tests were administered to all seventh grade pupils of one Indiana junior high school. Just before and one week after group test results interpretation, the pupils estimated their abilities in those areas sampled by the tests and their test scores. Means and variances of the estimates were tested for significance of change using appropriate statistical techniques. Null hypotheses predicted no significant differences for the various combinations of estimates

for the total group and several subgroups. The results of these statistical procedures appear in Chapter IV.

CHAPTER IV

FINDINGS OF THE STUDY

On the following pages of this chapter, the various statistical findings of the study are presented, analyzed, and summarized.

After taking the four sections of the Academic Promise Tests, each subject in the study contributed four scores: a preinterpretation estimate of his abilities; a preinterpretation estimate of his test scores; a postinterpretation estimate of his abilities, and a postinterpretation estimate of his test scores. Estimates were made in terms of stanines, and in order to eliminate negative numbers, a constant of +32 was added. If, for example, an individual estimated all four of his test scores in the same stanine band as the scores actually fell, his estimate score was 32.00. Scores larger than 32 indicate over-estimations; scores smaller than 32, under-estimations.

I. STUDY GROUP NORMS COMPARED TO NATIONAL NORMS

Means and standard deviations of test scores for both the study group and the publisher's national norm group are presented in Table V, on page 47. The means and standard deviations of the study group were found to be very similar to those of the national norm group furnished by the test publisher.

II. MEANS OF ESTIMATES

Means of estimates appear in Table VI, page 48. These means were

TABLE V

ACADEMIC PROMISE TESTS NATIONAL NORMS COMPARED TO STUDY GROUP NORMS

	Means		Standard Deviations	
	Norm Group	Study Group	Norm Group	Study Group
Abstract Reasoning	28.7	31.27	12.2	13.3
Numerical	25.6	25.76	10.0	9.8
Verbal	29.0	33.36	9.5	9.3
Language Usage	27.8	28.45	11.0	9.9
A.P.T. Total	111.1	115.5	36.1	34.9
Boys	--	115.97	--	35.6
Girls	--	115.05	--	34.1

TABLE VI

MEANS OF ESTIMATES OF ABILITIES AND POSITIONS OF TEST RESULTS FOR BOYS, GIRLS, AND TOTAL GROUP

	Boys (125)		Girls (119)		Total (244)	
	Pre-	Post-	Pre-	Post-	Pre-	Post-
Estimate of abilities	33.94	34.45	35.40	35.31	34.65	34.87
Estimate of position of test results	33.56	33.37	34.36	33.79	33.95	33.57

tested for significance of difference, the results appearing in Table VII on page 50. No significant differences were found. This was not an unexpected finding, since test of significance of difference of means is not the most appropriate technique for this particular situation. An individual with two extreme over-estimates and two extreme under-estimates might have a total mean estimate not unlike the individual who estimated all four parts accurately. It seems more logical to look for changes in variability than in means.

III. VARIABILITY OF ESTIMATES FROM GRAND MEANS OF ESTIMATES

Table VIII, page 51, indicates the variances and standard deviations of the estimates of the study group, while Table IX on page 52 shows the results of testing the significance of variance of various combinations of preinterpretation and postinterpretation estimates by use of the F-test. After interpretation of test results, the total study group had less variance in both the estimates of abilities and the estimate of test results, at a highly significant level of confidence. For the total study group, no significant difference was found between the variances of the preinterpretation estimates of ability and the preinterpretation estimates of test results. After interpretation of test results, the total study group had less variance in the estimate of test results than in the estimate of abilities at a highly significant level of confidence.

Table X on page 53 indicates the variability of the boys' pre-interpretation and postinterpretation estimates of abilities and of test results as computed from the means of the total group, while Table XI on

TABLE VII

COMPARISON OF MEANS OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND OF MEANS OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS BY SEX AND BY TOTAL GROUP

	Pre Mean	Post Mean	D	$s_{\bar{X}-\bar{X}}$	t	Significance
Grand Means of Estimates of abilities	34.65	34.87	- .22	.46	-0.48	--
Grand Means of Estimates of Test Results	33.95	33.57	+ .38	.42	+0.91	--
Boys' Means of Estimates of Abilities	33.94	34.45	- .51	.60	-0.85	--
Boys' Means of Estimates of Test Results	33.56	33.37	+ .19	.31	+0.34	--
Girls' Means of Estimates of Abilities	35.40	35.31	+ .09	.69	+0.13	--
Girls' Means of Estimates of Test Results	34.36	33.79	- .57	.63	-0.91	--

TABLE VIII

VARIABILITY OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITY AND OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM GRAND MEANS OF THE RESPECTIVE ESTIMATES

	Σd		Σd^2				
	boys	girls	boys	girls	total	variance	standard deviation
Pre-estimate of abilities	-90.25	+89.65	3828.90	4645.88	8474.78	34.7327	5.8935
Pre-estimate of test results	-48.75	+48.95	3725.50	4173.30	7898.80	32.3721	5.6897
Post-estimate of abilities	-52.75	+56.47	1781.56	2108.04	3889.60	15.9410	3.9926
Post-estimate of test results	-29.25	+25.17	1111.98	1364.92	2476.90	10.1512	3.1861

TABLE IX

F-TESTS OF SIGNIFICANCE OF VARIANCE FOR VARIOUS COMBINATIONS OF TOTAL GROUP PREINTERPRETATION ESTIMATES AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND PREINTERPRETATION AND POST - INTERPRETATION ESTIMATES OF TEST RESULTS

Numerator - Denominator Comparisons	F	Significance
Pre-estimate of abilities with post-estimate of abilities	2.1788	.002
Pre-estimate of test results with post-estimates of test results	3.1889	.002
Pre-estimate of abilities with pre-estimates of test results	1.0729	---
Post-estimate of abilities with post-estimates of test results	1.5703	.002

TABLE X

VARIABILITY OF BOYS' PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITY AND PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM THE GRAND MEANS OF ESTIMATES

	Variance	Standard Deviation
Pre-estimate of abilities	30.6312	5.5345
Pre-estimate of test results	29.8040	5.4593
Post-estimate of abilities	14.2524	3.7752
Post-estimate of test results	8.8958	2.9826

TABLE XI

F-TESTS OF SIGNIFICANCE OF VARIANCE FOR BOYS' PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM GRAND MEANS OF ESTIMATES

Numerator-Denominator Combination	F	Significance
Pre-estimate of abilities with post-estimate of abilities	2.1492	.002
Pre-estimate of test results with post-estimate of test results	3.3503	.002
Pre-estimate of abilities with pre-estimate of test results	1.0278	---
Post-estimate of abilities with post-estimate of test results	1.6021	.002

the same page shows the results of testing the significance of the changes in variance. After test interpretation, the boys had less variance in both the estimates of abilities and in the estimates of test results, differences significant at a highly significant level of confidence. No significant difference was found between the variances of the preinterpretation estimates of abilities and the preinterpretation estimates of test results. However, after test interpretation, the boys had less variance in the estimates of test results than in the estimates of abilities, a difference significant at a high level of confidence.

Table XII on page 55 indicates the variability of the girls' preinterpretation and postinterpretation estimates of abilities and of test results as computed from the means of the total group, while Table XIII on the same page shows the results of testing the significance of the changes in variance. After test interpretation, the boys had less variance in both the estimates of abilities and in the estimates of test results, differences significant at a highly significant level of confidence. No significant difference was found between the variances of the preinterpretation estimates of abilities and the preinterpretation estimates of test results. However, after test interpretation, the girls had significantly less variance in the estimates of test results than in the estimates of abilities.

Table XIV on page 56 indicates the results of comparing variances from grand means of various combinations of boys' and girls' preinterpretation and postinterpretation estimates of abilities and of test results. Although the boys had a tendency toward less variance in both the pre-

TABLE XII

VARIABILITY OF GIRLS' PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITY AND PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM THE GRAND MEANS OF ESTIMATES

	Variance	Standard Deviation
Pre-estimate of abilities	39.0410	6.2483
Pre-estimate of test results	35.0697	5.9220
Post-estimate of abilities	17.7146	4.2089
Post-estimate of test results	11.4699	3.3867

TABLE XIII

F-TESTS OF SIGNIFICANCE OF VARIANCE FOR GIRLS' PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM GRAND MEANS OF ESTIMATES

Numerator-Denominator Combination	F	Significance
Pre-estimate of abilities with post-estimate of abilities	2.2038	.002
Pre-estimate of test results with post-estimate of test results	3.0575	.002
Pre-estimate of abilities with pre-estimate of test results	1.1132	---
Post-estimate of abilities with post-estimate of test results	1.5444	.02

TABLE XIV

F-TESTS OF SIGNIFICANCE OF VARIANCE FOR VARIOUS COMBINATIONS OF BOYS' AND GIRLS' PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM GRAND MEANS OF ESTIMATES

Numerator-Denominator Comparisons	F	Significance
Girls' pre-estimate of abilities with boys' pre-estimate of abilities	1.2745	.10
Girls' pre-estimate of test results with boys' pre-estimate of test results	1.1766	--
Girls' post-estimate of abilities with boys' post-estimate of abilities	1.2429	--
Girls' post-estimate of test results with boys' post-estimate of test results	1.2893	.10

interpretation estimate of abilities and the postinterpretation estimate of test results, the differences did not reach an acceptable level of confidence. No significant differences between the sexes were found.

IV. VARIABILITY OF ESTIMATES FROM SEX MEANS

Table XV on page 58 indicates the variability of the boys' preinterpretation and postinterpretation estimates of abilities and of test results as computed from the means of the sex group, while Table XVI on the same page shows the results of testing the significance of the changes in variance. After test interpretation, the boys had less variance in both the estimates of abilities and in the estimates of test results, differences significant at a highly significant level of confidence. No significant difference was found between the variances of the preinterpretation estimates of abilities and the preinterpretation estimates of test results. However, after test interpretation, the boys had significantly less variance in the estimates of test results than in the estimates of abilities.

Table XVII on page 59 indicates the variability of the girls' preinterpretation and postinterpretation estimates of abilities and of test results as computed from the means of the sex group, while Table XVIII on the same page shows the results of testing the significance of the changes in variance. After test interpretation, the girls had less variance in both the estimates of abilities and in the estimates of test results, differences significant at a highly significant level of confidence. No significant difference was found between the variances of

TABLE XV

VARIABILITY OF BOYS' PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITY AND PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM THE MEANS OF BOYS' ESTIMATES

	Σd	Σd^2	Variance	Standard Deviation
Pre-estimate of abilities	- .50	3765.04	30.1203	5.4882
Pre-estimate of test results	0.00	3720.83	29.7666	5.4559
Post-estimate of abilities	-.25	1758.60	14.0688	3.7508
Post-estimate of test results	-.25	1140.20	9.1216	3.0202

TABLE XVI

F-TESTS OF SIGNIFICANCE OF VARIANCE FOR BOYS' PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM MEANS OF BOYS' ESTIMATES

Numerator-Denominator Combination	F	Significance
Pre-estimate of abilities with post-estimate of abilities	2.1409	.002
Pre-estimate of test results with post-estimate of test results	3.2633	.002
Pre-estimate of abilities with pre-estimate of test results	1.0119	--
Post-estimate of abilities with post-estimate of test results	1.5424	.02

TABLE XVII

VARIABILITY OF GIRLS' PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITY AND PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM THE MEANS OF GIRLS' ESTIMATES

	Σd	Σd^2	Variance	Standard Deviation
Pre-estimate of abilities	-9.60	4578.64	38.4760	6.2029
Pre-estimate of test results	+ .16	3988.38	33.5158	5.7893
Post-estimate of abilities	+ .11	2241.96	18.8400	4.3405
Post-estimate of test results	- .01	1358.86	11.4190	3.3792

TABLE XVIII

F-TESTS OF SIGNIFICANCE OF VARIANCE FOR GIRLS' PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM MEANS OF GIRLS' ESTIMATES

Numerator-Denominator Combination	F	Significance
Pre-estimate of abilities with post-estimate of abilities	2.0423	.002
Pre-estimate of test results with post-estimate of test results	2.9351	.002
Pre-estimate of abilities with pre-estimate of test results	1.1480	--
Post-estimate of abilities with post-estimate of test results	1.6499	.002

the preinterpretation estimates of abilities and the preinterpretation estimates of test results. However, after test interpretation, the girls had less variance in the estimates of test results than in the estimates of abilities, a difference significant at a highly significant level of confidence.

Table XIX on page 61 indicates the results of comparing variances from the appropriate sex means of various combinations of boys' and girls' preinterpretation and postinterpretation estimates of abilities and of test results. Although the boys tended toward less variance in all four estimates, the differences did not reach an acceptable level of confidence.

V. MEANS OF THE HIGH AND LOW ACADEMIC APTITUDE GROUPS

Subjects were divided into high, middle, and low academic aptitude groups on the basis of their total A.P.T. score in order to investigate the relationship of academic aptitude to accuracy of estimation. Those whose scores fell at or above .6 standard deviation above the mean of the study group were put into the high academic aptitude group, and those whose score fell at or below .6 standard deviation below this mean were placed in the low academic aptitude group. Table XX on page 62 indicates the results of comparing the means of these groups for significance of difference. After test interpretation, both the mean estimate of abilities and the mean estimate of test results of the high academic aptitude group were significantly higher. After test interpretation, both the mean estimate of abilities and the mean estimate of test results of the low academic aptitude group were significantly lower.

TABLE XIX

F-TESTS OF SIGNIFICANCE OF VARIANCE FOR VARIOUS COMBINATIONS OF BOYS' AND GIRLS' PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM THE MEANS OF BOYS' ESTIMATES AND THE MEANS OF THE GIRLS' ESTIMATES

Numerator-Denominator Comparisons	F	Significance
Girls' pre-estimate of abilities with boys' pre-estimate of abilities	1.2774	.10
Girls' pre-estimate of test results with boys' pre-estimate of test results	1.1259	.10
Girls' post-estimate of abilities with boys' post-estimate of abilities	1.3391	.10
Girls' post-estimate of test results with boys' post-estimate of test results	1.2518	.10

TABLE XX

COMPARISON OF MEANS OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND OF MEANS OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS OF THE HIGH ACADEMIC APTITUDE GROUP AND THE LOW ACADEMIC APTITUDE GROUP*

	Preinter-pretation Mean	Postinter-pretation Mean	D	$s_{\bar{X}-\bar{X}}$	t	Significance
High Academic Group estimates of abilities	29.91	32.77	- 2.86	.8950	- 3.1955	.01
High Academic Group estimates of test results	29.14	32.36	- 3.22	.8544	- 3.7687	.01
Low Academic Group estimates of abilities	39.22	36.50	+ 2.72	1.145	+ 2.3756	.05
Low Academic Group estimates of test results	37.86	34.91	+ 2.95	.8953	+ 3.2950	.01

*The high academic aptitude group is composed of those subjects whose Total A.P.T. scores fell at or above .6 standard deviation from the mean of the total A.P.T. scores of the total study group. The low academic aptitude group is composed of those subjects whose Total A.P.T. scores fell at or below .6 standard deviation from the mean of total A.P.T. scores of the total study group.

VI. MEANS OF THE OVER-ESTIMATORS AND UNDER-ESTIMATORS OF ABILITY

Subjects were also divided into three groups on the basis of their degree of accuracy in initial estimates of their abilities. Those whose preinterpretation estimates of ability fell at or above .6 standard deviation above the mean of estimates of ability of the total study group formed the over-estimators of ability group. In a similar fashion, those whose preinterpretation estimates of ability fell at or below .6 standard deviation below the mean of the preinterpretation estimates of abilities composed the under-estimators of ability group. Table XXI on page 64 reports the results of testing the means of estimates of these groups for significance of change. After test interpretation, both the mean estimate of abilities and the mean estimate of test results of the over-estimators of ability group were significantly lower. After test interpretation, both the mean estimate of abilities and the mean estimate of test results of the under-estimators of ability group were significantly higher.

VII. MEANS OF THE OVER-ESTIMATORS AND UNDER-ESTIMATORS OF TEST RESULTS

Finally, over-estimators and under-estimators of test results were identified on the basis of their initial estimates of test results. Those whose preinterpretation estimates of test results fell at or above .6 standard deviation above the mean of preinterpretation estimates of test results of the total group were placed in the over-estimators of test results group. Those whose preinterpretation estimates of test

TABLE XXI

COMPARISONS OF MEANS OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND MEANS OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS OF THE OVER-ESTIMATORS OF ABILITY AND OF THE UNDER-ESTIMATORS OF ABILITY*

	Preinterpretation Mean	Postinterpretation Mean	D	$s_{\bar{X}-\bar{X}}$	t	Significance
Over-estimators of ability, estimates of ability	41.41	37.76	+ 3.65	1.0665	+ 3.4224	.01
Over-estimators of ability, estimates of test results	37.49	35.53	+ 1.96	.9403	+ 2.0844	.05
Under-estimators of ability, estimates of ability	27.42	31.71	- 4.29	1.0042	- 4.2721	.01
Under-estimators of ability, estimates of test results	27.97	32.30	- 4.33	.8901	- 4.8646	.01

*The over-estimators of ability group is composed of those subjects whose pre-estimates of ability fell at or above .6 standard deviation from the mean of the pre-estimates of ability of the total study group. The under-estimators of ability group is composed of those subjects whose pre-estimates of ability fell at or below .6 standard deviation from the mean of the pre-estimates of ability of the total study group.

results fell at or below .6 standard deviation below the mean of estimates of test results formed in the under-estimators of test results group. Table XXII on page 66 contains the results of testing the means of estimates of these groups for significance of change. No significant difference was found between the mean of the preinterpretation estimate of abilities and the mean of the postinterpretation estimate of abilities of the over-estimators of ability group. However, after test interpretation the mean estimate of test results of the over-estimators of test results group was significantly lower. For the under-estimators of ability group, both the postinterpretation estimates of abilities and the postinterpretation estimates of test results were significantly higher.

VIII. MEANS OF THE SUB-GROUPS COMPARED TO GRAND MEANS OF THE TOTAL STUDY GROUP

The next step in the analysis of the data was to compare the means of the above-mentioned sub-groups with the means of the total study group. Table XXIII on page 67 shows the comparison of the means of the high academic aptitude group with the means of the total group. All four mean estimates of the high academic aptitude group were significantly lower than were the mean estimates of the total group in the study. For the high academic aptitude group, the mean postinterpretation estimates of abilities and of test results were very close to the points of maximum accuracy.

Table XXIV on page 67 shows the comparison of the means of the low academic aptitude group with the means of the total study group.

TABLE XXII

COMPARISONS OF MEANS OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS OF THE OVER-ESTIMATORS OF TEST RESULTS AND OF THE UNDER-ESTIMATORS OF TEST RESULTS*

	Preinter-pretation Mean	Postinter-pretation Mean	D	$s_{\bar{X}-\bar{X}}$	t	Significance
Over-estimators of test results, estimates of ability	40.06	38.17	+ 1.89	.9625	+ 1.9636	--
Over-estimators of test results, estimates of test results	40.28	35.75	+ 4.53	.9119	+ 4.9677	.01
Under-estimators of test results, estimates of abilities	29.15	31.66	- 2.51	.8766	- 2.8633	.01
Under-estimators of test results, estimates of test results	27.71	32.01	- 4.30	.8233	- 5.2229	.01

*The over-estimators of test results group is composed of those subjects whose pre-estimates of test results fell at or above .6 standard deviation from the mean of the pre-estimates of test results of the total study group. The under-estimators of test results group is composed of those subjects whose pre-estimates of test results score fell at or below .6 standard deviation from the mean of the total study group.

TABLE XXIII

COMPARISON OF THE MEANS OF THE HIGH ACADEMIC GROUP WITH THE MEANS OF THE TOTAL STUDY GROUP

	Total Group in the Study	High Academic Aptitude Group	D	$s_{\bar{X}-\bar{X}}$	t	Significance
Pre-estimate of abilities	34.65	29.91	+ 4.74	.8710	+ 5.442	.001
Pre-estimate of test results	33.95	29.14	+ 4.81	.8923	+ 5.3906	.001
Post-estimate of abilities	34.87	32.77	+ 2.10	.5001	+ 4.1992	.001
Post-estimate of test results	33.57	32.36	+ 1.21	.3289	+ 3.6790	.001

TABLE XXIV

COMPARISON OF THE MEANS OF THE LOW ACADEMIC GROUP WITH THE MEANS OF THE TOTAL STUDY GROUP

	Total Group in the Study	Low Academic Aptitude Group	D	$s_{\bar{X}-\bar{X}}$	t	Significance
Pre-estimate of abilities	34.65	39.22	- 4.57	.8925	- 5.1204	.001
Pre-estimate of test results	33.95	37.86	- 3.91	.8115	- 4.8182	.001
Post-estimate of abilities	34.87	36.50	- 1.63	.6288	- 2.5922	.01
Post-estimate of test results	33.57	34.91	- 1.34	.5617	- 2.3856	.05

All four mean estimates of the low academic aptitude group were significantly higher and farther from the point of maximum accuracy than were the mean estimates of the total group in the study.

Table XXV on page 69 shows the comparison of the means of the over-estimators of ability group with the means of the total study group. All four mean estimates of the over-estimators of ability group were significantly higher and farther from the point of maximum accuracy than were the mean estimates of the total group in the study.

Table XXVI on page 69 shows the comparison of the means of the under-estimators of ability group with the means of the study group. All four mean estimates of the under-estimators of ability group were significantly lower than were the mean estimates of the total group in the study. The mean postinterpretation estimates of abilities and of test results of the under-estimators of ability group were very close to the points of maximum accuracy.

Table XXVII on page 70 shows the comparison of the means of the over-estimators of test results with the means of the total study group. All four mean estimates of the over-estimators of test results group were significantly higher and farther from the point of maximum accuracy than were the mean estimates of the total group in the study.

Finally, Table XXVIII on page 70 shows the comparison of the means of the under-estimators of test results group with the means of the total study group. All four mean estimates of the under-estimators of test results group were significantly lower than were the mean estimates of the total study group. For the under-estimators of test results group,

TABLE XXV

COMPARISON OF THE MEANS OF THE OVER-ESTIMATORS OF ABILITY GROUP WITH THE MEANS OF THE TOTAL STUDY GROUP

	Total Group in the Study	Over-estimators of ability group	D	$s_{\bar{X}-\bar{X}}$	t	Significance
Pre-estimate of abilities	34.65	41.41	- 6.76	.9635	- 7.0161	.001
Pre-estimate of test results	33.95	37.49	- 3.54	.8768	- 4.0374	.001
Post-estimate of abilities	34.87	37.76	- 2.89	.6455	- 4.4771	.001
Post-estimate of test results	33.57	35.53	- 1.96	.5681	- 3.450	.001

TABLE XXVI

COMPARISON OF THE MEANS OF THE UNDER-ESTIMATORS OF ABILITY GROUP WITH THE MEANS OF THE TOTAL STUDY GROUP

	Total Group in the Study	Under-estimators of ability group	D	$s_{\bar{X}-\bar{X}}$	t	Significance
Pre-estimate of abilities	34.65	27.42	+ 7.23	.9461	+ 7.6419	.001
Pre-estimate of test results	33.95	27.97	+ 5.98	.9193	+ 6.5049	.001
Post-estimate of abilities	34.87	31.71	+ 3.16	.5668	+ 5.5752	.001
Post-estimate of test results	33.57	32.30	+ 1.27	.3468	+ 3.6620	.001

TABLE XVII

COMPARISON OF THE MEANS OF THE OVER-ESTIMATORS OF TEST RESULTS GROUP WITH THE MEANS OF THE TOTAL STUDY GROUP

	Total Group in the Study	Over-estimators of Ability Group	D	$s_{\bar{X}-\bar{X}}$	t	Significance
Pre-estimate of abilities	34.65	40.06	- 5.41	.8690	- 6.2255	.001
Pre-estimate of test results	33.95	40.28	- 6.33	.8634	- 7.3314	.001
Post-estimate of abilities	34.87	38.17	- 3.30	.6157	- 5.3598	.001
Post-estimate of test results	33.57	35.75	- 2.18	.5116	- 4.2611	.001

TABLE XVIII

COMPARISON OF THE MEANS OF THE UNDER-ESTIMATORS OF TEST RESULTS GROUP WITH THE MEANS OF THE TOTAL STUDY GROUP

	Total Group in the Study	Under-estimators of Test Results Group	D	$s_{\bar{X}-\bar{X}}$	t	Significance
Pre-estimate of abilities	34.65	29.15	+ 5.50	.8321	6.6010	.001
Pre-estimate of test results	33.95	27.71	+ 6.24	.8026	7.7747	.001
Post-estimate of abilities	34.87	31.66	+ 3.21	.5328	6.025	.001
Post-estimate of test results	33.57	32.01	+ 1.56	.3457	4.5126	.001

the mean postinterpretation estimates of abilities and of test results were very close to the points of maximum accuracy.

IX. PROPORTIONS OF BOYS AND GIRLS IN THE TOTAL STUDY GROUP AND IN THE VARIOUS SUB-GROUPS

The total study group and each of the six sub-groups were examined to determine the significance, if any, of the proportions of boys and girls. This data appears in Table XXIX on page 72. No group in the study, including the total group, was found to differ significantly in the proportion of boys and girls, from what would ordinarily be expected.

X. VARIABILITY OF ESTIMATES OF THE HIGH AND LOW ACADEMIC APTITUDE GROUPS FROM THE GRAND MEANS OF ESTIMATES

As noted previously, the major statistical technique in the study was to be the analysis of variance of the preinterpretation and post-interpretation estimates of abilities and of test results. Table XXX on page 73 presents the variances and standard deviations of the four estimates of the high academic aptitude group computed from the total group's means. Table XXXI on the same page contains the results of testing these variances for significance of change. After test interpretation, there was a highly significant decrease in the variances of both the estimates of abilities and the estimates of test results for the high academic aptitude group. No significant difference in variance was found between the preinterpretation estimate of test results and the preinter-pretation estimate of abilities. After interpretation of test results there was less variance in the estimates of test results than in the

TABLE XXIX

CHI SQUARE TESTS OF SIGNIFICANCE OF THE PROPORTIONS OF BOYS AND GIRLS IN THE VARIOUS GROUPS OF THE STUDY

Total Group in the Study		observed	expected	X^2	Significance
	boys	125	122		
	girls	119	122		
		244	244	.1474	--
High Academic Aptitude Group		**observed**	**expected**	**X^2**	**Significance**
	boys	31	34		
	girls	35	32		
		66	66	.5459	--
Low Academic Aptitude Group		**observed**	**expected**	**X^2**	**Significance**
	boys	38	40		
	girls	40	38		
		78	78	.2052	--
Over-Estimators of Ability		**observed**	**expected**	**X^2**	**Significance**
	boys	35	38		
	girls	39	36		
		74	74	.4868	--
Under-Estimators of Ability		**observed**	**expected**	**X^2**	**Significance**
	boys	39	36		
	girls	32	35		
		79	79	.5017	--
Over-Estimators of Test Results		**observed**	**expected**	**X^2**	**Significance**
	boys	37	40		
	girls	42	39		
		79	79	.4557	--
Under-Estimators of Test Results		**observed**	**expected**	**X^2**	**Significance**
	boys	40	41		
	girls	40	39		
		80	80	.0499	--

TABLE XXX

VARIABILITY OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITY AND OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM GRAND MEANS OF THE RESPECTIVE ESTIMATES FOR THE HIGH ACADEMIC APTITUDE GROUP

	Σd	Σd^2	Variance	Standard Deviation
Pre-estimate of abilities	- 312.90	2684.72	40.6776	6.3779
Pre-estimate of test results	- 314.70	2890.20	43.7910	6.6175
Post-estimate of abilities	- 134.42	805.18	12.1997	3.4928
Post-estimate of test results	- 46.62	290.24	4.3976	2.0970

TABLE XXXI

F-TESTS OF SIGNIFICANCE OF VARIANCE FOR VARIOUS COMBINATIONS OF HIGH ACADEMIC APTITUDE GROUPS PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM GRAND MEANS OF ESTIMATES

Numerator-Denominator Comparisons	F	Significance
Pre-estimate of abilities with post-estimate of abilities	3.3343	.002
Pre-estimate of test results with post-estimate of test results	9.9579	.002
Pre-estimate of test results with pre-estimate of abilities	1.0765	--
Post-estimate of abilities with post-estimate of test results	2.7741	.002

estimates of abilities, a difference significant at a highly significant level of confidence.

Table XXXII of page 75 presents the variances and standard deviations of the four estimates of the low academic aptitude group, while Table XXXIII on the same page reports the results of testing these variances for significance of difference. After test interpretation, the low academic aptitude group had a highly significant decrease in variances of estimates of both abilities and of test results. No significant differences in variance were found between either the pre-interpretation estimates of abilities and the preinterpretation estimates of test results, or between the postinterpretation estimates of abilities and the postinterpretation estimates of test results.

XI. VARIABILITY OF ESTIMATES OF THE OVER-ESTIMATORS OF ABILITY AND OF THE UNDER-ESTIMATORS OF ABILITY FROM THE GRAND MEANS OF ESTIMATES

Table XXXIV of page 76 presents the variances and standard deviations of the four estimates of the over-estimators of ability group, while Table XXXV on the same page reports the results of testing these variances for significance of difference. After test interpretation there was a highly significant decrease in the variances of both the estimates of abilities and the estimates of test results. No significant differences in variances were found between either the preinterpretation estimate of abilities and the preinterpretation estimate of test results, or between the postinterpretation estimate of abilities and the postinterpretation estimate of test results.

TABLE XXXII

VARIABILITY OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITY AND OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM THE GRAND MEANS OF THE RESPECTIVE ESTIMATES FOR THE LOW ACADEMIC APTITUDE GROUP

	Σd	Σd^2	Variance	Standard Deviation
Pre-estimate of abilities	+ 365.30	3980.66	51.0341	7.1438
Pre-estimate of test results	+ 304.90	3199.10	41.0141	6.4042
Post-estimate of abilities	+ 137.14	2008.08	25.7446	5.0739
Post-estimate of test results	+ 100.54	1666.66	21.3674	4.6225

TABLE XXXIII

F-TESTS OF SIGNIFICANCE OF VARIANCE FOR VARIOUS COMBINATIONS OF LOW ACADEMIC APTITUDE GROUP'S PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM GRAND MEANS OF ESTIMATES

Numerator-Denominator Combinations	F	Significance
Pre-estimate of abilities with post-estimate of abilities	1.9823	.002
Pre-estimate of test results with post-estimate of test results	1.9194	.002
Pre-estimate of abilities with pre-estimate of test results	1.2443	--
Post-estimate of abilities with post-estimate of test results	1.0976	--

TABLE XXXIV

VARIABILITY OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITY AND OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM THE GRAND MEANS OF THE RESPECTIVE ESTIMATES FOR THE OVER-ESTIMATORS OF ABILITY GROUP

	Σd	Σd^2	Variance	Standard Deviation
Pre-estimate of abilities	+ 499.90	4304.71	58.1718	7.6270
Pre-estimate of test results	+ 411.70	3482.80	47.0649	6.8604
Post-estimate of abilities	+ 213.52	1924.52	26.0070	5.0997
Post-estimate of test results	+ 144.82	1539.50	20.8041	4.5612

TABLE XXXV

F-TESTS OF SIGNIFICANCE OF VARIANCE FOR VARIOUS COMBINATIONS OF OVER-ESTIMATORS OF ABILITY GROUP'S PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM GRAND MEANS OF ESTIMATES

Numerator-Denominator Combinations	F	Significance
Pre-estimate of abilities with post-estimate of abilities	2.2367	.002
Pre-estimate of test results with post-estimate of test results	2.2622	.002
Pre-estimate of abilities with pre-estimate of test results	1.2359	--
Post-estimate of abilities with post-estimate of test results	1.2500	--

Table XXXVI of page 78 presents the variances and standard deviations of the four estimates of the under-estimators of ability group, while Table XXXVII on the same page reports the results of testing these variances for significance of difference. After test interpretation, the under-estimators of ability group had a highly significant decrease in variances of both the estimate of abilities and of test results. No significant difference in variance was found between the preinterpretation estimate of abilities and the preinterpretation estimate of test results for this group. However, after test interpretation, the under-estimators of ability group had less variance in their estimates of test results than in their estimates of abilities, a difference significant at a highly significant level of confidence.

XII. VARIABILITY OF ESTIMATES OF THE OVER-ESTIMATORS OF TEST RESULTS AND OF THE UNDER-ESTIMATORS OF TEST RESULTS FROM THE GRAND MEANS OF ESTIMATES

Table XXXVIII of page 79 presents the variances and standard deviations of the four estimates of the over-estimators of test results group, while Table XXXIX on the same page reports the results of testing these variances for significance of difference. After test interpretation, there was a highly significant decrease in the variances of both the estimates of abilities and the estimates of test results. No significant difference in variance was found between the preinterpretation estimate of abilities and the preinterpretation estimate of test results. For the over-estimators of test results group, there was a tendency in the direction toward more variance in the postinterpretation estimate of abilities than in the postinterpretation estimate of test results, but this difference did not reach the acceptable level of confidence.

TABLE XXXVI

VARIABILITY OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITY AND OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM THE GRAND MEANS OF THE RESPECTIVE ESTIMATES FOR THE UNDER-ESTIMATORS OF ABILITY GROUP

	Σd	Σd^2	Variance	Standard Deviation
Pre-estimate of abilities	- 483.15	3794.92	53.4496	7.3109
Pre-estimate of test results	- 424.45	3591.20	50.5803	7.1120
Post-estimate of abilities	- 223.77	1289.84	18.1688	4.2623
Post-estimate of test results	- 90.47	403.72	5.5862	2.3846

TABLE XXXVII

F-TESTS OF SIGNIFICANCE OF VARIANCE OF VARIOUS COMBINATIONS OF UNDER-ESTIMATORS OF ABILITY GROUP'S PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM GRAND MEANS OF ESTIMATES

Numerator-Denominator Combinations	F	Significance
Pre-estimate of abilities with post-estimate of abilities	2.9422	.002
Pre-estimate of test results with post-estimate of test results	8.8952	.002
Pre-estimate of abilities with pre-estimate of test results	1.0567	--
Post-estimate of abilities with post-estimate of test results	3.1948	.002

TABLE XXXVIII

VARIABILITY OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITY AND OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM THE GRAND MEANS OF THE RESPECTIVE ESTIMATES FOR THE OVER-ESTIMATORS OF TEST RESULTS GROUP

	Σd	Σd^2	Variance	Standard Deviation
Pre-estimate of abilities	+ 431.65	3824.28	48.4086	6.9576
Pre-estimate of test results	+ 499.95	3823.64	48.4005	6.9571
Post-estimate of abilities	+ 260.27	1958.58	24.7922	4.9792
Post-estimate of test results	+ 171.97	1367.46	17.3096	4.1605

TABLE XXXIX

F-TESTS OF SIGNIFICANCE OF VARIANCE FOR VARIOUS COMBINATIONS OF OVER-ESTIMATORS OF TEST RESULTS GROUP'S PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM GRAND MEANS OF ESTIMATES

Numerator-Denominator Combinations	F	Significance
Pre-estimate of abilities with post-estimate of abilities	1.9525	.002
Pre-estimate of test results with post-estimate of test results	2.7961	.002
Pre-estimate of abilities with pre-estimate of test results	1.0001	--
Post-estimate of abilities with post-estimate of test results	1.4322	.10

Finally, Table XXXX of page 81 presents the variances and standard deviations of the four estimates of the under-estimators of test results group, while Table XXXXI on the same page reports the results of testing these variances for significance of difference. After test interpretation, there was a highly significant decrease in the variances of both the estimates of abilities and the estimates of test results. No significant difference in variance was found between the preinterpretation estimate of test results and the preinterpretation estimate of abilities. After test interpretation, the under-estimators of test results group had less variance in their estimates of test results than in their estimates of abilities, a difference significant at a highly significant level of confidence.

XIII. VARIABILITY OF ESTIMATES OF HIGH ACADEMIC APTITUDE PUPILS COMPARED TO THE VARIABILITY OF ESTIMATES OF LOW ACADEMIC APTITUDE PUPILS

Table XXXXII on page 82 shows the results of comparing the variances of the four estimates of the high academic aptitude group with those from the low academic aptitude group. No significant differences in variances were found between either the preinterpretation estimates of abilities or the preinterpretation estimates of test results of these two sub-groups. However, after test interpretation the high academic aptitude group had less variance in both the estimate of abilities and the estimate of test results, differences significant at a highly significant level of confidence.

XIV. VARIABILITY OF ESTIMATES OF OVER-ESTIMATORS OF ABILITIES COMPARED TO THE VARIABILITY OF ESTIMATES OF UNDER-ESTIMATORS OF ABILITIES

Table XXXXIII on page 84 shows the results of comparing the

TABLE XXXX

VARIABILITY OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITY AND OF PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM THE GRAND MEANS OF THE RESPECTIVE ESTIMATES FOR THE UNDER-ESTIMATORS OF TEST RESULTS GROUP

	Σd	Σd^2	Variance	Standard Deviation
Pre-estimate of abilities	- 440.00	3520.03	44.0004	6.6333
Pre-estimate of test results	- 499.00	3912.70	48.9088	6.9288
Post-estimate of abilities	- 256.60	1399.18	17.4898	4.1821
Post-estimate of test results	- 124.60	498.52	6.2315	2.4963

TABLE XXXXI

F-TESTS OF SIGNIFICANCE OF VARIANCE FOR VARIOUS COMBINATIONS OF UNDER-ESTIMATORS OF TEST RESULTS GROUP'S PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS FROM GRAND MEANS OF ESTIMATES

Numerator-Denominator Combinations	F	Significance
Pre-estimate of abilities with post-estimate of abilities	2.5157	.002
Pre-estimate of test results with post-estimate of test results	7.8486	.002
Pre-estimate of test results with pre-estimate of abilities	1.1115	--
Post-estimate of abilities with post-estimate of test results	2.8066	.002

TABLE XXXXII

F-TESTS OF SIGNIFICANCE OF VARIANCE FOR PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES FOR PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS COMPARING HIGH ACADEMIC APTITUDE PUPILS WITH LOW ACADEMIC APTITUDE PUPILS

Numerator-Denominator Combinations	F	Significance
Low academic aptitude group pre-estimate of abilities with high academic aptitude group pre-estimate of abilities	1.2545	--
High academic aptitude group pre-estimate of test results with low academic aptitude group pre-estimate of test results	1.0677	--
Low academic aptitude group post-estimate of abilities with high academic aptitude post-estimate of abilities	2.1102	.002
Low academic aptitude group post-estimate of test results with high academic aptitude group post-estimate of test results	4.8588	.002

variances of the four estimates of the over-estimators of ability group with those from the under-estimators of ability group. No significant differences in variances were found between either the preinterpretation estimates of abilities or the preinterpretation estimates of test results of these groups. Although the under-estimators of abilities group had a smaller variance in the post-estimate of abilities, the difference did not reach the acceptable level of confidence. After test interpretation, the under-estimators of ability group had less variance in the estimates of test results than did the over-estimators of ability group, a difference significant at a highly significant level of confidence.

XV. VARIABILITY OF ESTIMATES OF OVER-ESTIMATORS OF TEST RESULTS COMPARED TO THE VARIABILITY OF UNDER-ESTIMATORS OF TEST RESULTS

Table XXXXIV on page 85 shows the results of comparing the variances of the four estimates of the over-estimators of test results group with those from the under-estimators of test results group. No significant differences in variances were found between either the preinterpretation estimates of abilities or the preinterpretation estimates of test results of these two groups. Although the under-estimators of test results group had less variance in the postinterpretation estimate of abilities than did the over-estimators of test results group, the difference did not reach the acceptable level of confidence. After test interpretation, the under-estimators of test results group had less variance in their estimate

TABLE XXXXIII

F-TESTS OF SIGNIFICANCE OF VARIANCE FOR PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND FOR PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS COMPARING THE OVER-ESTIMATORS OF ABILITY WITH THE UNDER-ESTIMATORS OF ABILITY

Numerator-Denominator Combinations	F	Significance
Over-estimators of ability group pre-estimate of abilities with under-estimators of ability group pre-estimate of abilities	1.0884	--
Under-estimators of ability group pre-estimate of test results with over-estimators of ability group pre-estimate of test results	1.0746	--
Over-estimators of ability group post-estimate of abilities with under-estimators of ability group post-estimate of abilities	1.4315	.10
Over-estimators of ability group post-estimate of test results with under-estimators of ability group post-estimate of test results	3.6587	.002

TABLE XXXXIV

F-TESTS OF SIGNIFICANCE OF VARIANCE FOR PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF ABILITIES AND FOR PREINTERPRETATION AND POSTINTERPRETATION ESTIMATES OF TEST RESULTS COMPARING OVER-ESTIMATORS OF TEST RESULTS WITH THE UNDER-ESTIMATORS OF TEST RESULTS

Numerator-Denominator Combinations	F	Significance
Over-estimators of test results group pre-estimate of abilities with under-estimators of test results group pre-estimate of abilities	1.0092	--
Under-estimators of test results group pre-estimate of test results with over-estimators of test results group pre-estimate of test results	1.0105	--
Over-estimators of test results group post-estimate of abilities with under-estimators of test results group post-estimate of abilities	1.4175	.10
Over-estimators of test results group post-estimate of test results with under-estimators of test results group post-estimate of test results	2.7777	.002

of test results than did the over-estimators of test results group, a difference significant at a highly significant level of confidence.

XVI. SUMMARY OF THE FINDINGS

The major findings of this study are summarized on the following pages. A considerable number of significant findings were obtained from the analysis of the data.

1. For the following groups, the post-interpretation means of estimates of abilities were significantly closer to the points of maximum accuracy of estimation than the preinterpretation 1means: the high and low academic aptitude groups; and over-estimators and under-estimators of abilities; and the under-estimators of test results. No significant changes in means toward the points of maximum accuracy were found for the total group, the boys, the girls, or the over-estimators of test results.

2. For the following groups, the postinterpretation means of estimates of test results were significantly closer to the points of maximum accuracy of estimation than the preinterpretation means: the high and low academic aptitude groups; the over-estimators and under-estimators of abilities and the over-estimators and under-estimators of test results. No significant changes in means toward the points of maximum accuracy were found for the total group, the boys, or the girls.

3. After test results interpretation there were highly significant decreases in the variances of the estimates of abilities for the total group as well as for every sub-group in the study.

4. After test results interpretation there were highly significant decreases in the variances of the estimates of test results for the total

group as well as for every sub-group in the study.

5. No significant differences in accuracy of estimation were found between preinterpretation estimates of abilities and preinterpretation estimates of test results.

6. For the following groups, the variances of the postinterpretation estimates of test results were significantly less than the variances of the postinterpretation estimates of abilities: the total study group; the boys and girls, regardless of whether the variances were computed from the grand means or the appropriate sex mean; the high academic aptitude group; the under-estimators of abilities; and the under-estimators of test results. (The differences for the total group, the boys when computed from the grand means, the girls when computed from the appropriate sex means, the high academic group, the under-estimators of abilities, and the under-estimators of test results were highly significant.) No significant differences between the postinterpretation estimates of abilities and postinterpretation estimates of test results were found for either the low academic aptitude group, the over-estimators of abilities, or the over-estimators of test results.

7. When variances of the boys and girls from grand means of estimates in the study were compared, it was found that there was a tendency toward less variance in both the preinterpretation estimates of abilities and in the postinterpretation estimates of test results for the boys, but the differences did not reach an acceptable level of confidence. Neither were significant differences in variability found in the preinterpretation estimates of test results or the postinterpretation estimates of abilities.

8. When the variances from the sex means of estimates of the boys and girls were compared, it was found that although the boys had a tendency toward smaller variances in all four estimates, the differences did not reach an acceptable level of confidence.

9. Neither the total study group nor any of the various sub-groups were found to differ significantly in the proportion of boys and girls.

10. When the mean estimates of the high academic aptitude group were compared with the mean estimates of the total study group, it was found that all four mean estimates of the high academic aptitude group were significantly lower than the mean estimates of the total study group. In addition, the postinterpretation estimates of the high academic aptitude group were very close to the points of maximum accuracy of estimation.

11. When the mean estimates of the low academic aptitude group were compared to the means of the total study group, it was found that all four mean estimates of the low academic aptitude group were significantly higher and farther from the points of maximum accuracy of estimation than were the mean estimates of the total group in the study.

12. When the mean estimates of the over-estimators of abilities group were compared with the means of the total study group, it was found that all four estimates of this sub-group were significantly higher and farther from the points of maximum accuracy than were the mean estimates of the total study group.

13. When the mean estimates of the under-estimators of abilities

group were compared with the means of the total study group, it was found that all four mean estimates of this sub-group were significantly lower than were the mean estimates of the total study group. In addition, the postinterpretation estimates of the under-estimators of abilities group were very close to the points of maximum accuracy of estimation.

14. When the mean estimates of the over-estimators of test results were compared with the means of the total study group, it was found that all four mean estimates of this sub-group were significantly higher and farther from the points of maximum accuracy than were the mean estimates of the total study group.

15. When the mean estimates of the under-estimators of test results group were compared with the means of the total study group, it was found that all four mean estimates were significantly lower than were the mean estimates of the total study group. In addition, the postinterpretation estimates of the under-estimators of test results group were very close to the points of maximum accuracy of estimation.

16. When the variances of the high academic aptitude group were compared with the variances of the low academic aptitude group, it was found that there were no significant differences in the variances of either the preinterpretation estimates of abilities or of the preinterpretation estimates of test results. However, the high academic aptitude group had significantly less variance in both the postinterpretation estimates of abilities and the postinterpretation estimates of test results.

17. When the variances of the over-estimators of abilities group were compared with the variances of the under-estimators of abilities group, it was found that there were no significant differences in the

variances of either the preinterpretation estimates of abilities or of the preinterpretation estimates of test results. Although there was a tendency toward less variance in the postinterpretation estimates of abilities for the under-estimators of abilities group, the difference did not reach an acceptable level of confidence. However, the under-estimators of abilities group had significantly less variance in the postinterpretation estimates of test results than did the over-estimators of abilities group.

18. When the variances of the over-estimators of test results group were compared with the variances of the under-estimators of test results group, it was found that there were no significant differences in the variances of either the preinterpretation estimates of abilities or of the preinterpretation estimates of test results. Although there was a tendency toward less variance in the postinterpretation estimate of abilities for the under-estimators of test results group, the difference did not reach an acceptable level of confidence. However, the under-estimators of test results group had significantly less variance in the postinterpretation estimates of test results group had significantly less variance in the postinterpretation estimates of test results than did the over-estimators of test results group.

CHAPTER V

SUMMARY, CONCLUSIONS, RECOMMENDATIONS

I. SUMMARY

The purposes of this study were to develop an approach to the communication and interpretation of test scores, and to use the procedure with a group of seventh grade pupils in order to investigate its effectiveness.

The Academic Promise Tests, an aptitude battery measuring abstract reasoning ability, numerical ability, verbal ability, and language usage ability, were administered to 281 seventh grade pupils of one urban junior high school in Indiana. One month later, after a short discussion of individual differences, abilities, and the concept of norm groups, the subjects estimated both their abilities in the four areas sampled by the tests as well as their performance on the tests taken a month previously. Estimates were made on a chart composed of numbers and cartoon symbols representing 100 boys and girls arranged in such a manner as to form a stanine distribution. Immediately after this estimation procedure, group interpretation of test results followed, using the test publisher's printed profile. These procedures took place in 10 different groups averaging 28.1 pupils each. Exactly one week later, the pupils again estimated their abilities and test results in the same manner as before.

Data for the study were collected from the 244 pupils (119 girls and 125 boys) who took the tests and made both preinterpretation and

postinterpretation estimates. Means and standard deviations were computed for the preinterpretation and postinterpretation estimates of both abilities and test results for the total study group and the following sub-groups: boys; girls; high academic aptitude pupils; low academic aptitude pupils; over-estimators of abilities; under-estimators of abilities; over-estimators of test results; and under-estimators of test results. The t-test technique was used to test the significance of difference between means, the F-test technique to test the significance of difference between variances, and the chi-square technique to test the significance of the proportions of boys and girls in the various groups.

The following is a brief summary of the major findings of the study. After test interpretation, for the total group as well as for every sub-group in the study, there were highly significant variance decreases in both estimates of abilities and estimates of test results. No significant difference in accuracy of estimation were found between the sexes. The interpretation procedure was especially effective for the high academic aptitude group, the under-estimators of abilities group, and the under-estimators of test results group.

II. CONCLUSIONS

From the analysis of the data in this study, the following conclusions seem to be justified.

1. The majority of the seventh grade pupils in this study changed their preinterpretation estimates of their abilities and their preinterpretation estimates of their test performance in the desired direction after participating in the test results interpretation procedure utilized

in this study. It is concluded from these data that if seventh grade pupils similar to the ones in this study receive interpretations of aptitude test results in a manner similar to the procedure used in this study, they can be expected to become more accurate in their estimates of abilities and test performance.

2. While no meaningful differences were found between the pre-interpretation estimates of abilities and the preinterpretation estimates of test results, the variances of the postinterpretation estimates of test results were generally found to be significantly less than the variances of the postinterpretation estimates of abilities. It is concluded from these data that junior high school age youngsters are more able to alter their estimates of their test performance than they are to alter their estimates of their abilities as a result of group test interpretation.

3. The experimental version of test results interpretation with groups of pupils was especially effective with the high academic aptitude group and the groups that under-estimated either their abilities or their test results. All three of these sub-groups had mean estimates of abilities and test results very close to the points of maximum accuracy of estimation as well as highly significant decreases in variability of these estimates after group test results interpretation. Therefore, it is concluded that a group test results interpretation procedure similar to the one utilized in this study is likely to be most effective with high academic aptitude pupils and those who under-estimate either their abilities or test performances.

4. The experimental version of test results interpretation with groups of pupils was somewhat less effective with the low academic aptitude group and the groups that over-estimated either their abilities or their test results. However, all three of these sub-groups had highly significant decreases in variability of estimates of both abilities and of test results after group test results interpretation. Therefore, it is concluded that a group test results interpretation procedure similar to the one utilized in this study is likely to be somewhat less effective with low academic aptitude pupils and those who over-estimate either their abilities or test performance.

5. No significant differences in performance was found between the boys and the girls in this study, and it is concluded that there is no meaningful difference in the capability of seventh grade boys and girls to profit from group test results interpretation as practiced in this study.

III. RECOMMENDATIONS FOR GROUP TEST RESULTS INTERPRETATION

On the basis of the results of this study and the conclusions formulated above, the following recommendations for group test results interpretation with junior high school pupils are warranted.

1. The results of this study have several implications for practical application. It now seems less reasonable to omit interpretation of test results because the available counselor time is insufficient for individual interpretations. For the majority of students, a group procedure such as the one used in this study is likely to be

an effective means of helping a student understand the meaning of his test scores.

2. It is recommended that before group test results interpretation, the interpreter secure from each pupil an estimate of how he perceives his relative abilities and how he feels he performed on the tests. On the basis of this study, it would seem beneficial to follow up the group interpretation with additional individual or small group counseling interviews with those students of low academic aptitude and those who seriously over-estimated either their abilities or their test performance. This recommendation is in addition to the general invitation that should be extended to any student who might want to discuss his test results further with a counselor.

3. The assumption that all pupils understand the concepts of human variability and individual differences seems unwarranted. It is recommended that the interpreter preface his interpretation of test results with a discussion of these concepts.

4. Although this study did not investigate directly the effect of visual aids, it is the investigator's opinion that the interpreter make maximum use of appropriate visual aids during the interpretation process.

IV. SUGGESTIONS FOR FURTHER RESEARCH

The need for further study of the procedures and practices of group test results interpretation seems apparent. The following are some suggestions for studies that might very well be undertaken by future investigators.

1. As an initial recommendation, this study should be replicated with other seventh grade pupils as well as with students in other grade levels. Such replications should be facilitated by the forms and procedures recorded in the appendix of this dissertation.

2. In this study, the various estimates of the subtests of the Academic Promise Tests were summed, producing one gross estimation score. A future investigation might analyze the data further to study possible differences in estimating the specific abilities and test score estimates.

3. As noted previously, only one week intervened between the date of interpretation and the follow-up collection of data. A future investigation might increase or decrease this intervening period in order to study the effect of time on the ability to integrate and recall such information.

4. This study used an aptitude test. Other similar studies might utilize other tests, such as achievement tests, in order to study any possible differences in estimation accuracy related to the type of trait being estimated.

5. In a future study, the interpretations of test scores might be done by the pupils' regular classroom or homeroom teachers. It is assumed that such interpretations would be preceeded by in-service training in test interpretation methods and procedures.

6. The contributions of various types of visual aids to the effectiveness of the interpretation process needs to be studied in detail. Future research is needed to shed light upon which segments of the interpretation process can be made more effective and efficient by the use of visual aids, as well as what form these visual aids should take.

7. Interpretation groups in this study were composed of pupils assigned to regular classroom sections of the seventh grade. No attempt was made to group pupils homogeneously either on the basis of accuracy of estimations or on the basis of test scores. These factors might be investigated in order to study the possible effects of the degree of homogeneity or heterogeneity of the interpretation group on the effectiveness of the interpretation process.

BIBLIOGRAPHY

BIBLIOGRAPHY

"Are Americans Over-Testing?" Overview, August, 1961.

Arnold, Dwight L. "Student Reaction to the Kuder," Personnel and Guidance Journal, September, 1958.

Association for Supervision and Curriculum Development, National Education Association, Perceiving, Behaving, Becoming. Washington, D. C.: N.E.A., 1962.

Barry, Ruth, and Beverly Wolf, Epitaph for Vocational Guidance. New York: Bureau of Publications, Teachers College, Columbia University, 1962.

Bennett, George K., Marjorie K. Bennett, Dorothy M. Clendenen, Jerome E. Doppelt, James H. Ricks, Jr., Harold G. Seashore, and Alexander G. Wesman. Academic Promise Tests Manual, 1962. New York: The Psychological Corporation, 1962.

Berg, Irwin A., "Test Score Interpretation and Client Confusions," Personnel and Guidance Journal, May 1956.

Bordin, E.S. "A Theory of Vocational Interests as Dynamic Phenomena," Educational and Psychological Measurement, 3, 1943.

Combs, Arthur W. "A Perceptual View of the Adequate Personality," Perceiving, Behaving, Becoming. Washington, D. C.: N.E.A., 1962.

Combs, Arthur and D. Snygg. Individual Behavior, Revised Edition. New York: Harper & Brothers, 1959.

Dressel, Paul L. and Ross W. Matteson. "The Effect of Client Participation in Test Interpretation," Educational and Psychological Measurement. X, 4, Winter, 1950.

Durost, W. N. "Present Progress and Needed Improvements in School Evaluation Programs," Educational and Psychological Measurements, 14.

Erikson, E. H. Childhood and Society. (New York: Norton, 1955).

Friedenberg, Edgar Z. The Vanishing Adolescent. New York: Dell Publishing Company, 1962.

Froehlich, Clifford P. "Must Counseling Be Individual?" Educational and Psychological Measurement, XVIII, 4, 1958.

Froehlich, C. P. and W. E. Mosher. "Do Counselees Remember Test Scores?" Journal of Counseling Psychology. I, Fall, 1954.

Gibson, Robert L. "Pupil Opinions of High School Guidance Programs," Personnel and Guidance Journal, XL, 5, January, 1962.

Goldman, Leo. Using Tests in Counseling. New York: Appleton-Century-Crofts, 1961.

Gustad, John W. and Abdul H. Tuma. "The Effects of Different Methods of Test Introduction and Interpretation on Client Learning in Counseling," Journal of Counseling Psychology, Vol. 4, No. 4, 1957.

Hardaway, Charles W., Grace B. Kozak, and Helen Ederle. Budgetary and Financial Provisions for Guidance Services in Indiana Public Secondary Schools. Office of Research and Testing, Indiana State College, Terre Haute, Indiana, 1963.

Hastings, et al. "The Use of Test Results - Cooperative Research Project Number 509." Bureau of Educational Research, University of Illinois, 1960.

Jersild, Arthur T. In Search of Self. New York: Bureau of Publications, Teachers College, Columbia University, 1952.

Kamm, Robert B. and Gilbert C. Wrenn. "Client Acceptance of Self-Information in Counseling," Educational and Psychological Measurement, X, 1, Spring, 1950.

Lallas, John E. "A Comparison of Three Methods of Interpretation of Achievement Tests To Pupils," (unpublished doctoral dissertation, Sanford University, 1956).

Lane, David. "A Comparison of Two Techniques of Interpreting Test Results to Clients in Vocational Counseling," (unpublished doctoral dissertation, Columbia University, 1952).

McCarthy, Dorothea. "Ethical and Professional Considerations in Reporting Test Information," Proceedings and Summaries, 23d Annual Meeting, New York State Psychological Association, Inc., May 5-7, 1960.

McNemar, Quinn. Psychological Statistics. New York: John Wiley & Sons, Inc., 1955.

1961 Annual Survey of Standardized Test Publishers, prepared by Stanley B. Hunt and Associates for the American Textbook Publishers Institute, 432 Park Avenue South, New York City, New York, April, 1962.

Peters, Herman J. and Gail F. Farwell. Guidance: A Developmental Approach. Chicago: Rand McNally & Company, 1959.

Phillips, Beeman N. and Garrett Weathers. "Analysis of Errors Made in Scoring Standardized Tests," Educational and Psychological Measurement, 18, 1958.

Rogers, Lyle B. "A Comparison of Two Kinds of Test Interpretation Interview," Journal of Counseling Psychology, 1954.

Singer, Stanley L. and Buford Stefflre. "Analysis of the Self-Estimate in the Evaluation of Counseling," Journal of Counseling Psychology, I, No. 4, Winter, 1954.

Super, Donald E. "Vocational Adjustment: Implementing a Self-Concept," Occupations, XXX, November, 1951.

Super, Donald E. The Psychology of Careers. New York: Harper & Brothers, 1959.

"Testing: Can Everyone Be Pigeonholed?" Newsweek, 54, 3, July 20, 1959.

"Testing is Big Business," American Psychologist, II, January, 1947.

Thorndike, E. L. "Tests and Their Uses," Teachers College Record, XXVI, 2, October, 1924.

Thorndike, Robert L. and Elizabeth Hagen. Measurement and Evaluation in Psychology and Education. Second Edition. New York: Wiley & Sons, 1961.

Tyler, Leona E. The Work of the Counselor, Second Edition. New York: Appleton-Century-Crofts, 1961.

White, Robert W. The Abnormal Personality. New York: The Ronald Press, 1956.

Wright, Edward Wayne. "A Comparison of Individual and Multiple Counseling in the Dissemination and Interpretation of Test Data" (unpublished doctoral dissertation, University of California, 1957).

APPENDIX A

TEST INTERPRETATION SCRIPT

Today we are going to spend some time talking about abilities of seventh graders. Perhaps a good place to start would be to point out that most of the time we learn about our abilities by comparing our performances with the performances of others. For example, many seventh grade boys like to build models, and most seventh grade girls like to dance. You probably have noticed that before you really know how good you are at building models or dancing, you have to compare your models or your dancing with the models and dancing of your friends. As another example, suppose you are told that on a quiz of 60 spelling words, you got 40 correct. We really can't say for sure whether or not this is a high score, an average score, or a low score until we have the answers to at least two questions. First, how hard are the words? (It would make a big difference if they came from a college science book or a third grade reader!) Second, what were the scores of the other seventh graders who took this quiz? (If you knew this, you could easily see if your score was among the highest, about in the middle, or among the lowest.)

This chart[1] is supposed to represent 100 seventh grade students of your age. I would like for you to imagine that one of these symbols represents <u>you</u>. Therefore, if you are a girl, imagine that this represents you, 49 girls, and 50 boys, all in the seventh grade. If you are a boy, imagine that this represents you, 49 other boys, and 50 girls, all in the seventh grade.

Notice that what we have here is an arrangement of 5 rows of

---

[1]Appendix B.

20 students each, alternating the boys and girls. This is one way to arrange our 100 students, but there are many others. For instance, if we were to arrange them in the order of some ability, we would probably end up with an arrangement that looks something like this. Here we have the same 100 students arranged in order of their relative amounts of some ability. Some would demonstrate a high amount of the ability and would rank about here; some would not have much of the ability and would rank about here; many would have an average amount of the ability and would rank about here; some would have an above-average amount of the ability and would rank about here; and some would have a below-average amount of the ability and would rank about here. In order that you might see this arrangement a little closer, I have prepared duplicates on individual sheets, which will be passed out now.[2]

At the same time we will pass out a mimeographed worksheet.[3] Please print your name and group number on the lines at the top. The schedule number for this group is ___.

Now place the distribution on top of your worksheet so that the first paragraph and box shows at the top.

Now, there are many, many abilities that we could discuss today, but we are going to consider only four. Brief descriptions of these abilities are printed on your worksheet. Let's take a look at Ability Number 1. Read the description silently as I read it aloud:

> Abstract Reasoning--the ability to see relationships and recognize concepts presented in the form of diagrams or symbols rather than in words or numbers.

---

[2]Appendix B.

[3]Appendix C.

Now: look at the distribution. Imagine that these 100 seventh grade students are arranged from bottom to top in order of their increasing amount of abstract reasoning ability. That is, those with a low amount of this ability here, then those with a below-average amount, then those with an average amount, then those with an above-average amount, and finally, those with the most of this ability. Remember that this distribution is supposed to represent 100 seventh graders, including you. Now, in such an arrangement, find the symbol which you feel best represents the position in this group that illustrates <u>your</u> ability in abstract reasoning, as compared with the others. (Pause) When you have found that symbol, look to the left and note the number of the row in which the symbol you have selected appears. Write that number in the first box on your paper.

OK, slide your distribution sheet down so we can take a look at Ability Number 2. Read the description silently as I read it aloud:

> Numerical--the capacity to think in quantitative terms, to understand and use numerical relationships.

Look at the distribution again. This time, imagine that these 100 seventh grade students are arranged from bottom to top in order of their increasing amount of numerical ability. That is, from those with the least amount of this ability at the bottom on up to those with the greatest amount at the top. Remember that you are to suppose that one of these 100 students represents you. In such an arrangement, find the symbol which you feel best represents the position in this group that illustrates your numerical ability, as compared with the others. (Pause) When you have located that symbol, look to the left to see the number of the row it is in. Write this number in the second box on your paper.

Now let's take a look at Ability Number Three on your paper. Read the description silently as I read it aloud:

> Verbal--understanding word meanings and the ability to use words in reasoning.

This time, imagine that your distribution represents 100 seventh graders arranged from bottom to top in order of their increasing abstract reasoningability. As before, those with the least are at the bottom, those with the most at the top. One of these symbols is supposed to represent you. Find the symbol which you feel best represents the position in this group that illustrates your verbal ability, as compared with the others, Put the number of the row in which your symbol appears in box number three.

Finally, let's consider Ability Number Four on your paper. Read it's description silently as I read it aloud:

> Language Usage--understanding of correct writing and speech, appreciation of good English.

You probably have guessed correctly that this time your distribution is supposed to represent 100 seventh graders, including yourself, arranged from bottom to top in order of an increasing amount of language usage ability. Find the symbol which you feel best represents the position in this group that illustrates your language usage ability, as compared to the others. When you have found it, write the number of the row in which it appears in box number four. Now cover your worksheet with your printed distribution.

From this point on, please do not erase or change any number you have written in the boxes. You will not be graded in any way on this, of course. The reason is that I am interested in knowing your first impression. OK?

Now, you have just estimated your relative amounts of these abilities. There are other ways to estimate relative ability. Do you remember this test you all took about a month ago? (Show.) Well, tests like these also estimate relative abilities. Do you remember this first part, the pages of puzzle-like problems? (Show.) Well, this is supposed to be one way of estimating your abstract reasoning ability.

Look at your copy of this distribution of 100 seventh grade students again. This time, imagine that these symbols are arranged in order of scores on this part of the test, from those who made the lowest scores here at the bottom to those who made the highest scores here at the top. If we were to arrange 100 seventh grade students who took this abstract reasoning test in this manner, and you were one of that 100, where do you think you would be assigned to stand? That is, which symbol do you think represents your performance on this part of the test as compared to others of your age and grade who took it? Find that symbol; locate the number of the row in which it appears, and copy that number on the dotted line next to the first box. Note: this number may or may not be the same as the one you have previously written in the box. If it is, OK; if it is larger, OK; if it is smaller, OK; this is up to you to decide.

Do you remember the second part of the test, the one with the arithmetic problems? This part of the test is supposed to estimate your numerical ability. Imagine this time that the 100 students in the distribution are arranged from bottom to top in order of scores on this part of the test, from those who made the lowest scores here at the bottom to those who made the highest scores here at the top. If we were to arrange 100 seventh grade students who took this test in this manner, you were

one of that 100, where do you think you would be assigned to stand? Which symbol do you think represents your performance on this numerical part of the test as compared to others of your age and grade who took it? After you have located that symbol, see what row it is in, and copy that number on the dotted line beside box number two. As before, this number may or may not be the same as the one you put in the box previously.

Remember the part of the test that was concerned with word meanings, and how words were related to each other? This part of the test attempts to estimate verbal ability. Imagine that your distribution now represents 100 seventh grade pupils arranged in order of their scores on the verbal part of the test, from those who made the lowest scores at the bottom to those who made the highest scores at the top. If you were one of the 100 students arranged in such a manner, where do you think you would be assigned to stand? Which symbol do you think represents your performance on this verbal part of the test as compared to others of your age and grade who took it? Locate that symbol; note the number of the row in which it appears, and write that number on the dotted line beside the third box. As before, this number may or may not be the same as the one you wrote on the box number three.

Finally, do you remember the last part of the test in which you were asked to inspect some sentences that were divided into three parts and tell whether there was an error of some kind in part A, part B, part C, or if there were no errors? This part of the test is supposed to estimate language usage ability. Look at your distribution and imagine this time that the 100 students are arranged from bottom to top in order of their increasing scores on the language usage part of the test. That is, with those who made the lowest scores at the bottom, and those who made the highest scores at the top. If

you were one of the 100 students arranged in this manner, where do you think you would be assigned to stand? Which symbol do you think represents your performance on this verbal part of the test, as compared to others of your age and grade who took it? Find your symbol; note the number of the row in which it appears, and write that number on the dotted line beside the last box. As before, this number may or may not be the same as the one you wrote in box number four.

Check your worksheet; you should have your name printed at the top, your group number, ____, in the proper space, one number in each of the boxes, and one number on each of the dotted lines. Now pass your worksheets to the front of the room and then the printed distributions also.

Now let's review what we have done so far. You have estimated your relative abilities in abstract reasoning, numerical, verbal, and language usage. Then you estimated your performance on the tests that measures these abilities. The next step will be to reveal to you your test results. Please keep your test profile sheet face down until everyone has theirs. Also, although you may not care if others see your test results, perhaps they might not want you to see theirs. For this reason, it will be best if you keep your eyes on your own paper.

Turn your papers over. The printing on the left describes what this test is supposed to measure, and how to use this profile sheet in understanding your scores. We are interested in the first four columns. Let's look at the first column, which tells us about your results on the abstract reasoning part of the test. Notice the two numbers above this first column. The one in the gray area does not concern us now, but the one in the white area, beside the word "percentile" does. A percentile tells what per cent

of the students of your age and grade who took this part of the test made scores at or below your score. Think of it as being the number of students out of 100 who made scores at or below yours. This number says to you, "Out of 100 students of your age and grade who took this abstract reasoning part of the test, your score is as high or higher than this many."

Now look at the column below the first percentile number. Note the heavy black line in the middle, the one with the number 50 beside it. A score at this point would have approximately 50% or half of the other scores above it, and 50% or half of the other scores below it. Part of your column has been blackened with a crayon; this blackened area indicates how far and in what direction your score varied from the middle of the distribution. If your percentile score at the top of the first column is greater than 50, the blackened area should run up from the 50 point to your score. If your score was less than 50, the blackened column should run down to your score. Check to see that the percentile number at the top is the same as the number at the end of the blackened column.

The second column reports your performance on the numerical part of the test. Find the percentile number at the top. As before a percentile tells what per cent made scores at or below your score. Think of it as being the number of students out of 100 who made scores like or below yours. This number says to you, "Out of 100 students of your age and grade who took this numerical part of the test, your score is as high or higher than this many."

Check the blackened area of the second column to see if it has been marked correctly. If your percentile score is larger than 50, the column should run up from the 50 point to your score. If your percentile score is less than 50, the column should run down from the 50 point. The number at

the end of this column should be the same as the percentile number at the top.

Column three shows your performance on the verbal section of the test. Find the percentile number at the top of the column, which is interpreted just like the other percentiles: it shows what per cent of other seventh grade students who took this part of the test made scores at or below your score. It indicates the number of students out of 100 who made scores like or below yours. Think of it as saying to you, "Out of 100 students of your age and grade who took the verbal part of the test, your score is as high or higher than this many."

Check the blackened area of this column to see if it has been colored in correctly. If your score is larger than 50, the column should run up from the 50 point to your score. If your score is less than 50, the column should run down from the 50 point to your score. As before, the number at the end of this blackened column should match your percentile score at the top.

Finally, column 4 indicates your performance on the language usage section of the test. The percentile number at the top shows what per cent of the other students like you made scores at or below your score. It can be thought of as the number of students out of 100 who made scores like or below yours. It says, "out of 100 students of your grade who took this language part of the test, your score is as high or higher than this many."

Check the column 4 blackened area to see if it has been marked correctly. If your percentile score is larger than 50, it should run up from the 50 point. If your score is less than 50, it should run down from the 50 point. In either case the number at the side of the end of this column should be the same as the percentile number at the top.

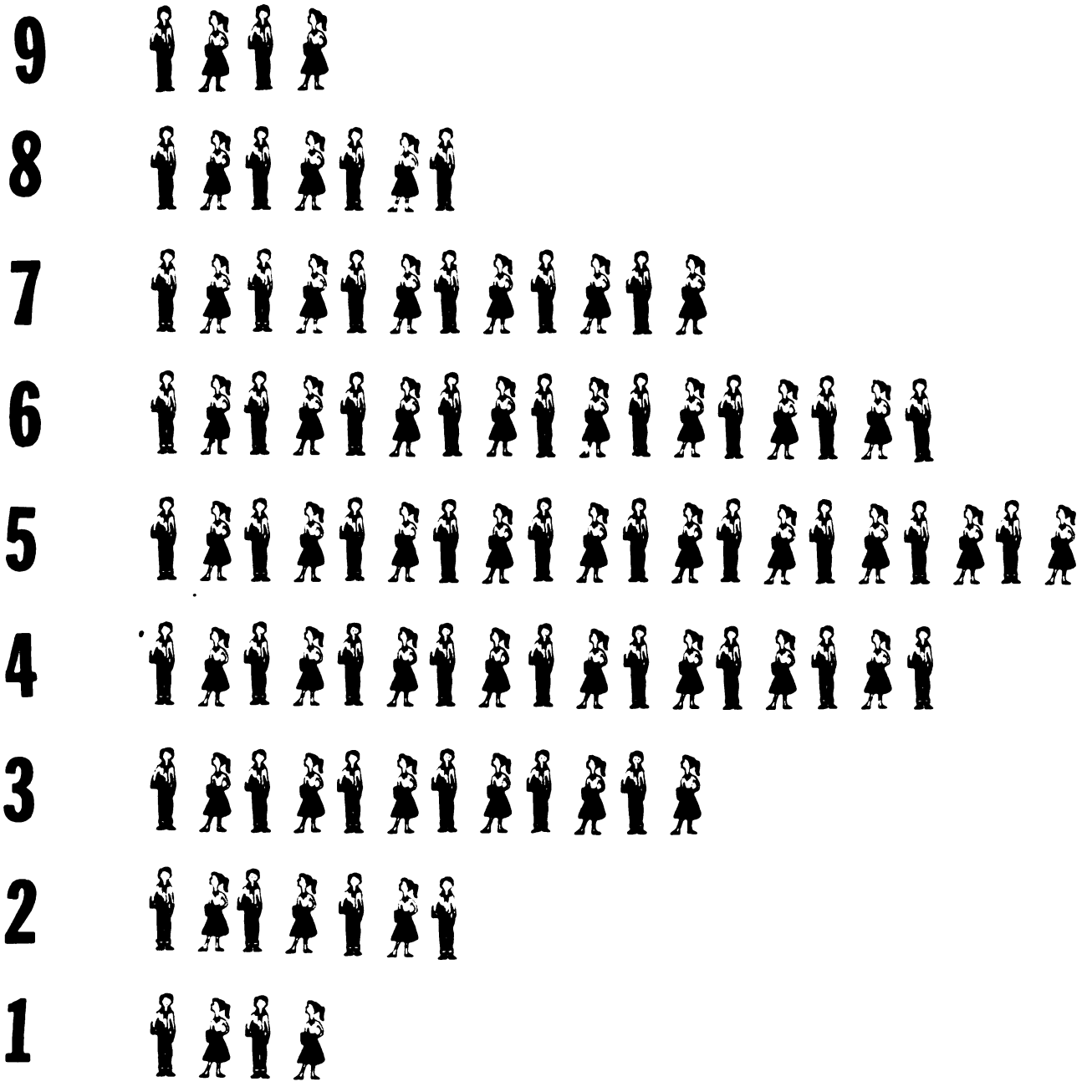
One thing further might be said about percentile scores. In general, scores that fall between 30 and 70 can be said to be about "average," those between 70 and 85 "above average," those between 15 and 30 "below average," those above 85 "high," and those below 15 "low." These are rough approximations, but are close enough for our purposes here.

Now who has a question I can try to answer?

In just a minute or so, we will take up your profiles. In a few weeks I will return them to Mr. Dobson in the Guidance Office on the second floor. Perhaps you would like to arrange an individual interview to discuss your test results. If so, he has asked me to tell you that he will be very happy to make an appointment with any student who would like to have one.

APPENDIX B

NORM GROUP CHART IN STANINE FORM


9
8
7
6
5
4
3
2
1

APPENDIX C

ESTIMATION WORKSHEET

Your name:_____Group number:_____
(Last) (Middle) (First)

Ability No. 1: Abstract Reasoning

...the ability to see relationships and recognize concepts presented in the form of diagrams or symbols rather than in words or numbers ..... [ ]

Ability No. 2: Numerical

...the capacity to think in quantitative terms, to understand and use numerical relationships ..... [ ]

Ability No. 3. Verbal

...understanding word meanings and the ability to use words in reasoning ..... [ ]

Ability No. 4. Language Usage

...understanding of correct writing and speech, appreciation of good English ..... [ ]

APPENDIX D

FOLLOW-UP SCRIPT

Today we would like for you to take a few minutes to make a few more estimates for us. First we will pass out the materials, the printed distribution chart[1] and a worksheet.[2] (Pause) Print your name and group number in the indicated spaces at the top of the worksheet. The schedule-number of this group is _____.

Perhaps you were absent on the day of the test a month ago, or perhaps you were absent a week ago when I was here before. If so, write, "I was absent on the day of the test," or "I was absent a week ago" above your name on the worksheet.

One caution: please do not jump to conclusions as to what we will be doing today! These directions are a little bit different than they were last week. Listen carefully until each section of the directions has been completed so you won't make a mistake and have to erase.

Let me say at this point that no one in this school will see your responses on this worksheet, so I hope you will try to be as accurate in your estimations as possible. And of course you will be in no way graded on your responses; I am just interested in learning what you think. Because this is a matter between you and me, it will be best if you keep your paper covered.

Let's review briefly what this distribution is supposed to represent. Remember that we pointed out that if we arranged 100 individuals in order of their increasing amounts of some ability, we would probably get a distribution similar to this one. Those with the least amount of the ability

---

[1]Appendix B.

[2]Appendix C.

would rank about here, those with a little more about here, those with an average amount in this region, those with an above-average amount about here, and those with the most of the ability here. For convenience the rows have been numbered.

The first thing we would like for you to do today is to indicate your understanding of your test scores. First, place the distribution sheet on top of your worksheet in such a manner that all that shows of the worksheet are the dotted lines on the right and the group number at the top.

Look at the distribution. Imagine that these are 100 seventh grade students arranged from bottom to top in order of their scores on the first section of the test you took, the abstract reasoning section. One of these symbols is supposed to represent you. Find the symbol that is located in the position that indicates your test score in abstract reasoning as compared to the others. Look to the left, note the number of the row in which this symbol appears, and write that number on the first dotted line.

Next, imagine that these 100 seventh grade students are arranged from bottom to top in order of their scores on the second section of the test you took, the numerical section. As before, one of the symbols is supposed to represent you. Which symbol is standing in a position that represents your numerical test score as compared to the others? Look to the left, note the number of the row in which this symbol appears and write that number on the second dotted line.

Now, imagine that these 100 seventh grade students are arranged from bottom to top in order of their scores on the third section of the test, the verbal section. One of them is supposed to be you. Which symbol is

located in the position that indicates your test score on the verbal section of the test as compared to the others? When you have found that symbol, look to the left, note the number of the row in which it appears, and write that number on the third dotted line.

Finally, imagine now that the 100 seventh grade students are arranged from bottom to top in order of their scores on the last part of the test, the language usage section. As usual, one of these symbols is supposed to represent you. Find the symbol that is located in the position that indicates your test score in language usage as compared to the others. Look to the left, note number of the row in which your symbol appears, and write that number on the last dotted line.

Now, cover your worksheet with your distribution sheet.

Please pay careful attention to the next directions. Remember that we pointed out last time that tests are only one way of estimating abilities. There are many other ways, including making your own estimates. This is what we would like for you to do next, to indicate your own estimations of your relative abilities.

Place the distribution sheet on top of the worksheet so that the dotted lines and the group number space is covered. On the left you should see the descriptions of four abilities, each followed by a box.

Read the description of the first ability silently as I read it aloud:

> Abstract Reasoning--the ability to see the relationships and recognize concepts presented in the form of diagrams or symbols rather than in words or numbers.

Look at the distribution. Imagine that you see 100 seventh grade students arranged from bottom to top in order of increasing abstract reasoning

ability. One of them represents you. How much abstract reasoning ability do you think you have in comparison with the others? Find the symbol which is located in a position that you feel accurately indicates your abstract reasoning ability as compared with the others in the distribution. (Pause) Note the number of the row in which it appears, and write that number in the first box. Please note: this number may or may not be the same as the one you wrote on the first dotted line. If you feel that it should be larger, OK, if you feel it should be smaller, OK, and if you feel that it should be the same, OK. We are interested in knowing what you think, how you estimate your abstract reasoning ability.

Read the description of the second ability silently as I read it aloud:

> Numerical--the capacity to think in quantitative terms, to understand and use numerical relationships.

Look at your distribution. Imagine that you see 100 seventh grade students arranged from bottom to top in order of increasing numerical ability. One of them represents you. How much numerical ability do you think you have in comparison with the others? Find the symbol which is located in a position that you feel accurately indicates your numerical ability as compared with the others in the distribution. (Pause) Note the number of the row in which it appears, and write that number in the second box. As before, it may or may not be the same as the number you wrote on the second dotted line. We are interested in knowing what you think, how you estimate your numerical ability.

Finally, read the description of the third ability as I read it aloud:

Verbal--understanding word meanings and the ability to use words in reasoning.

Look at the distribution, and this time imagine that the 100 seventh grade students are arranged from bottom to top in order of increasing verbal ability, one of them representing you. How much verbal ability do you think you have in comparison with the others? Find the symbol which is located in a position that you feel accurately indicates your numerical ability as compared with the others in the distribution. (Pause) Note the number of the row in which it appears, and write that number in the third box. As before, it may or may not be the same as the number you wrote on the third dotted line. We are interested in knowing what you think, how you estimate your verbal ability.

Finally, read the description of the fourth ability silently as I read it aloud:

Language Usage--understanding of correct writing and speech, appreciation of good English.

Look at your distribution, and imagine that the 100 seventh grade students are arranged from bottom to top in order of increasing language usage ability, one of them representing you. How much verbal ability do you think you have in comparison with the others? Find the symbol which is located in a position that you feel accurately indicates your numerical ability as compared with the others in the distribution. (Pause) Note the number in the last box. As usual, it may or may not be the same number as you wrote on the last dotted line. We are interested in knowing what you think, how you estimate your language ability.

Take a look at your paper--you should have your name and group number, _____, at the top, a number from 1 to 9 on each of the dotted

lines, and a number from 1 to 9 in each of the boxes.

Pass your worksheets to the front of the room, face down. Then pass in the distribution sheets.