THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

A FACTOR ANALYTICAL STUDY OF HIGH SCHOOL PLACEMENT TEST, BATTERY FOR ENTERING FRESHMEN (FORM A)

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF EDUCATION

BY

ROBERT GEORGE DODSON

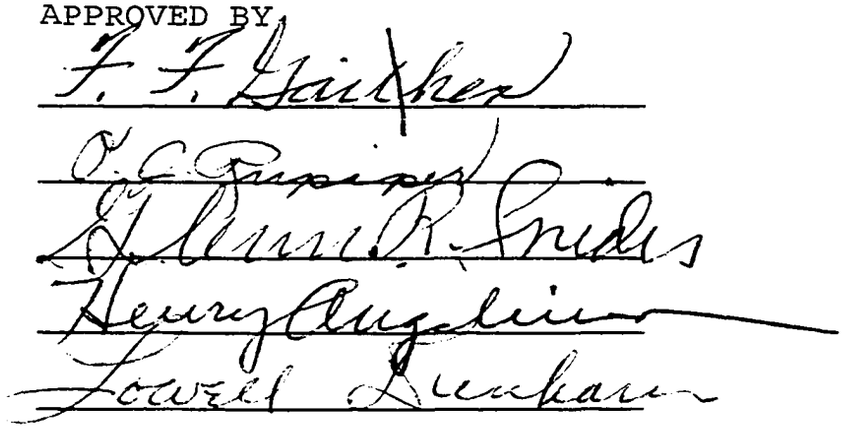
Norman, Oklahoma

1964

A FACTOR ANALYTICAL STUDY OF HIGH SCHOOL PLACEMENT TEST,

BATTERY FOR ENTERING FRESHMEN (FORM A)

APPROVED BY



DISSERTATION COMMITTEE

ACKNOWLEDGEMENTS

The author is especially grateful to Dr. Funston F. Gaither, chairman of the doctoral committee and major professor, for his constructive criticism and encouragement throughout the graduate program and for his patience and guidance during the time in which this report was developed.

Acknowledgement of profound indebtedness is due Dr. Omer J. Rupiper, director of this study, for the assistance, guidance, and time which he so generously gave in the development of the study and of this report.

The author is also indebted to Dr. Henry Angelino, Dr. Glenn R. Snider, and Dr. Lowell Dunham, members of the doctoral committee, for their suggestions, encouragement, and assistance during the graduate program and for their comments on the manuscript.

An expression of gratitude is also due Dr. Stephen Scatori, a member of the original doctoral committee removed by an untimely death, who befriended the author some twenty years ago and whose scholarship and urbanity continue to in-

spire and influence the life and professional development of the author.

A final acknowledgement of indebtedness is due the members of the staff of the Computer Laboratory of the University of Oklahoma for their assistance in the statistical analysis of the data used in this study and to Mrs. Els Gyermek for typing the manuscript to form.

TABLE OF CONTENTS

TABLE OF CONTENTS-Continued

LIST OF TABLES

LIST OF ILLUSTRATIONS

A FACTOR ANALYTICAL STUDY OF HIGH SCHOOL PLACEMENT TEST, BATTERY FOR ENTERING FRESHMEN (FORM A)

CHAPTER I

INTRODUCTION

Background of the Problem

The administration of high school placement tests for entering high school students has long been an important feature of a good secondary school guidance program. Such tests are utilized together with the pre-high school cumulative record of the student for guidance and counseling, for placement, and for sectioning of students in the high school curriculum. While the pattern of the content of such tests varies from school to school, there is substantial agreement that, at least, the placement tests should comprise a test of scholastic aptitude and achievement tests in reading, mathematics, and English.

Early in the development of testing programs as broad as that envisioned here, the problem of the possible lack of

comparability of the data derived from tests designed by different test experts and standardized on different normative groups demonstrated the desirability of designing what have come to be known as high school placement batteries. Today, batteries which have been standardized on the same samples and designed by the same team of experts have greatly reduced the problem of comparability. The battery generally consists of an instrument for determining mental ability or Mental Age and of tests for the measurement of attainment in reading, mathematics, and English.

While test authors had constructed several good high school placement tests for secondary schools in general, only in 1955 did such a test intended for the specific purpose of utilization in Catholic schools appear. This battery, the High School Placement Test, Battery for Entering Freshmen, (Form A), was constructed by Anderhalter, Colestock, and Gawkowski, and published by Scholastic Testing Service, Incorporated, of Chicago.[1] Form B was published in 1958. Since the appearance of Form B an alternate form has appeared each year. The publishers are now located in Bensenville, Illi-

---

[1]Oscar Krisen Buros (ed.), The Fifth Mental Measurements Yearbook (Highland Park, New Jersey: The Gryphon Press, 1959), p. 29.

nois. The features which made this battery popular among Catholic high school counselors were the inclusion of a test on religion as well as the fact that the normative sample included students completing Catholic elementary schools and applying for admission to Catholic high schools.

The High School Placement Test has enjoyed widespread utilization in Catholic schools throughout the country. Nevertheless, the battery appears to lack the scientific merits upon which such general acceptance ought to be founded. According to Cottle, there is "no way to determine whether the normative group is representative of all students entering the approximately 2,000 Catholic high schools in the United States . . . More evidence of concurrent validity needs to be provided."[1] Jones is more critical of the battery than Cottle. He states that "the marked limitations of the battery with respect to normative and validity data, the level of estimated reliabilities, and the inadequate control of the possible influence of certain types of response sets pose serious obstacles as to the overall utility of the instrument."[2]

---

[1]William C. Cottle and Robert A. Jones, "High School Placement Test," The Fifth Mental Measurements Yearbook, ed. Oscar Krisen Buros (Highland Park, New Jersey: The Gryphon Press, 1959), p. 29.

[2]Ibid.

Statement of the Problem

The problem investigated in this study was the determination of the adequacy of the High School Placement Test in general and of the religion subtest in particular. More specifically, answers to the following questions were sought: What abilities are measured by the battery? Does the battery meet the criterion of factorial validity? What is the nature of the religion subtest, that is, what does it measure and do the items meet the criteria of discrimination and difficulty level? The problem, therefore, consisted in the isolation and comparison of the factors yielded by the scores obtained on the tests included in this investigation and in an item analysis of the religion test.

Purpose of the Study

There is no battery of placement tests available to counselors in Catholic secondary schools which includes a test on religion and which is standardized in its entirety on the same population other than the High School Placement Test. This battery, however, appears to fail to meet the scientific criteria necessary to establish its adequacy. Nevertheless, counselors in Catholic high schools continue to use the battery extensively. These circumstances give rise to the need

to undertake a thorough scientific analysis of the battery. The purpose of this study, therefore, was the investigation of the adequacy of the battery and of the religion subtest. The investigation was implemented by attempting to identify the abilities measured by the battery and by attempting to establish the factorial validity of the subtests by isolating the factors unique to the various tests under study. Moreover, the religion subtest was given special consideration by subjecting it to an item analysis.

Limitations of the Study

The investigation was necessarily limited by a number of factors. In order to keep the data within manageable limits only four other standardized tests were employed in the analysis. These tests, together with the High School Placement Test and the first semester grade point average in ninth-grade religion, yielded a controlled total of twenty variables. The sample was relatively small (N = 59). Other restrictive factors characterized the sample in that all the constituents of the population were Catholic boys who were entering a Catholic college preparatory high school in the State of Oklahoma. Finally, the technique of multiple factor analysis itself does not purport to yield an exhaustive comprehension of the nature

of mental abilities. It achieves no more than yield data which are indicative of the use of one or more abilities in a given test performance by the specific population under consideration.

Procedure of the Study

The 20 variables used as the basis for the computation of the coefficients of correlation in this study were derived from the scores on the subtests of the five standardized tests and the first semester grade point average in ninth-grade religion obtained by the sample. The tests employed were the High School Placement Test, Battery for Entering Freshmen, (Form A),[1] the California Achievement Tests Complete Battery, Advanced, Grades 9 to 14, Form AA,[2] the California Short-Form Test of Mental Maturity, Secondary, Grades 9 to 13, 1957 S-Form, Form A,[3] the California Test of Personality, 1953 Re-

---

[1]O. F. Anderhalter, R. Colestock, and R. S. Gawkowski, High School Placement Test, Battery for Entering Freshmen, (Form A) (Chicago: Scholastic Testing Service, Inc., 1955).

[2]E. W. Tiegs and W. W. Clark, California Achievement Tests Complete Battery, Advanced, Grades 9 to 14, Form AA (Los Angeles: California Test Bureau, 1950).

[3]W. W. Clark and E. W. Tiegs, California Short-Form Test of Mental Maturity, Secondary, Grades 9 to 13, 1957 S-Form, Form A (Los Angeles: California Test Bureau, 1957).

vision, Secondary, Grades 9 to College, Form AA,[1] and the Religion Test for High Schools, Form B.[2] The tests were administered with strict adherence to the directions for administration found in the respective manuals. All answer sheets were machine scored by IBM equipment and checked by hand scoring. Zero-order coefficients of correlation of the obtained raw scores among the subtests were used to construct the matrix of intercorrelations which was factor analyzed. The religion test was further subjected to a complete item analysis.

Organization of the Study

This study is reported in six chapters and a reference section. The background of the problem, the statement of the problem, the purpose of the study, the limitations of the study, and the procedure of the study have been discussed in Chapter One. Chapter Two is devoted to a review of selected investigations related to the present study. The design of the study in terms of the subjects of the sample, the tests

---

[1]L. P. Thorpe, W. W. Clark, and E. W. Tiegs, California Test of Personality, 1953 Revision, Secondary, Grades 9 to College, Form AA (Los Angeles: California Test Bureau, 1953).

[2]A. G. Schmidt and O. F. Anderhalter, Religion Test for High Schools, Form B (Chicago: Loyola University Press, 1953).

used in the study, and the analytical procedure is discussed in Chapter Three. Chapter Four consists of the data which were collected and the interpretation of the findings by means of tables, a figure, and discussion. The item analysis of the religion subtest of the High School Placement Test is presented in Chapter Five. The summary of the investigation, the conclusions, and the recommendations are found in Chapter Six. The references used in this study are listed in the reference section.

CHAPTER II

REVIEW OF RELATED LITERATURE

Introduction

Many scientific investigations of the nature and structure of human abilities have been published in the past several decades. A review of the pertinent studies will contribute to a better understanding of the problem investigated in this study. Since the literature of factorial analysis is extremely extensive, the studies reviewed here are, of necessity, limited in number. Brief reference is made to the work of some of the researchers who have made substantial contributions to the development and popularization of factorial analysis, who have summarized and recapitulated the findings of others, or who have attempted to cross-identify the factors tentatively identified by others. More extensive reviews of selected investigations germane to this study are presented. The criteria for the inclusion of these specific studies in the present investigation were: 1) one or more of the tests used in this investigation were included,

2) Thurstone's method of multiple factor analysis was utilized, and 3) the studies were reported upon since 1950.

General Survey of Early Investigations

In 1951 and again in 1953, French made the following observation in the introductions to his two surveys of investigations of factorial analysis:

> Since each factor analysis can only explore a relatively small aspect of the mind, and since agreement between analyses is far from perfect, it is necessary to bring many of them together in order to arrive at a comprehensive picture.[1]

This observation explains the motivation of all the researchers who have dedicated much time to the investigation and reconsideration of factorial analyses reported by others.

In 1928 Dodd attempted to recapitulate the findings of all the investigations of the nature and structure of human abilities made prior to that year.[2] Wolfle continued this effort, bringing the recapitulation of reports up to 1940 and

---

[1]John W. French, The Description of Aptitude and Achievement Tests in Terms of Rotated Factors ("Psychometric Monographs," No. 5; Chicago: University of Chicago Press, 1951), p. 1.

John W. French, The Description of Personality Measurements in Terms of Rotated Factors (Princeton: Educational Testing Service, 1953), p. 6.

[2]S. C. Dodd, "The Theory of Factors," Psychological Review, XXXV (1928), 211-31; 261-79.

listing 530 investigations of this nature in his report.[1] In 1946 Cattell published an excellent analysis of factor and intercorrelation studies in the personality domain.[2] Vernon analyzed selected contributions made in England and the United States between 1935 and 1950.[3] Carroll and Schweiker summarized the factorial studies carried out from 1939 to 1951.[4] In 1951 French attempted to clarify the contribution of such investigations to the understanding of mental abilities by means of reviews of 69 factorial analyses involving tests of aptitude and achievement.[5] Directing his attention to the study of personality constructs, French published a study of the findings of 123 factorial analyses of personality measurements in 1953.[6]

---

[1]Dael Wolfle, Factor Analysis to 1940 ("Psychometric Monographs," No. 3; Chicago: University of Chicago Press, 1940).

[2]R. B. Cattell, Description and Measurement of Personality (Yonkers, New York: World Book Co., 1946).

[3]P. E. Vernon, The Structure of Human Abilities (London: Methuen and Co., Ltd., 1950).

[4]J. B. Carroll and R. F. Schweiker, "Factor Analysis in Educational Research," Review of Educational Research, XXI (1951), 368-88.

[5]French, The Description of Aptitude and Achievement Tests in Terms of Rotated Factors.

[6]French, The Description of Personality Measurements in Terms of Rotated Factors.

The principal conclusions of the foregoing investigators have had important implications for the entire field of factorial analysis. These implications are, to be sure, germane to the present study, also. While reference to such conclusions is made later when an attempt is made to interpret the nature of the extracted factors, they are reported here in summary form.

The problem of the nature and structure of human abilities has been investigated in depth by all of the forenamed investigators. They agree, generally, that human abilities appear to reflect the composite interaction of two broadly defined factors identified tentatively as cognitive and conative. Two general sub-factors which contribute to the complex interaction of the cognitive and conative factors are tentatively identified as motor or mechanical and personality. The cognitive factor appears to consist of two highly correlated factors tentatively identified as verbal and mathematical. Sub-factors, tentatively identified as spatial, memory, perceptual speed, and reasoning contribute to these two correlated factors. The nature of the conative factor has not been extensively investigated and, therefore, is still not clearly defined. As for the structure of these abilities, these researchers have rejected the theories known as "monarchic,"

"oligarchic," "anarchic," and "hierarchic." In regard to the problem of structure, French substantially summarizes the conclusions of these investigators when he states,

> The most parsimonious hypothesis for mental structure is that it contains a rather large number of frequently overlapping, causally unrelated factors, which vary greatly in comprehensiveness and which are produced by units of experience, by hereditary mechanisms, or by any combination of these two types of influences.[1]

Important inferences made by these investigators on the basis of their conclusions as to the nature and structure of human abilities include: 1) ability measures should have high intercorrelations, 2) attainment measures should correlate highly, 3) ability measures should correlate highly with measures of attainment, 4) verbal ability should correlate in a highly significant way with education or schooling, 5) grade point average should not correlate highly with standardized measures of ability and attainment since the former are never pure measures, 6) measures of personality should reflect a fairly large number of general factors and overlapping group factors, 7) measures of personality which have high coefficients of reliability and validity should have significant intercorrelations with measures of ability and attainment, and

---

[1]French, The Description of Aptitude and Achievement Tests in Terms of Rotated Factors, p. 13.

8) measures of personality which do not have high coefficients of reliability and validity should not correlate significantly with measures of ability and attainment.

Survey of Selected Recent Investigations

The 12 selected factor analyses reported on here more specifically were considered representative of investigations of this nature completed since 1950. Each one included, at least, one test which was also used in this study. The relevancy of the findings and conclusions of these investigations are considered in Chapter Four when an attempt is made to isolate and identify the factors operative in the performance of the sample on the variables included in this investigation.

In 1950 Spivey investigated the factor patterns of an achievement test, a personality test, an intelligence test, and an interest inventory.[1] The subjects of the study were 235 men and 165 women between the ages of 17 and 27 years, enrolled in the Division of Humanities at John Muir College, Pasadena, California. The population was grouped by sex and administered the following tests: Kuder Preference Record,

---

[1]Gordon M. Spivey, "The Relationship between Temperament and Achievement of a Selected Group of John Muir College Students" (unpublished Doctor's dissertation, University of Southern California, Los Angeles, 1950).

Primary Mental Abilities Test, Guilford's Inventory of Factors STDCR, and Progressive (now California) Achievement Test, Advanced. The subtests of the batteries yielded 24 variables for each sex for which the tetrachoric coefficients of correlation were computed separately for each sex. The 552 intercorrelations were arranged into two correlation matrices which were subjected to factor analysis by the centroid method. The resulting factor matrices were rotated into a pattern of positive manifold and simple structure.

The intercorrelations and the nine tentatively identified factors gave evidence of relationships between the variables of temperament and achievement. For the men the factors were tentatively labeled: general scholastic, scientific, personality, spatial reasoning, numerical, mechanical, and literary. For the women the factors were: general scholastic ability and "social butterfly." The dispersion of the projections on the other factors inhibited their identification. The Progressive Achievement Test variables projected only minimally significant relationships with the personality test for the men and slightly significant relationships for the women. The only highly significant projections of the achievement test contributed to the variance of the factors identified as general scholastic, scientific, spatial reasoning, numerical,

and literary.

In 1956 Rupiper investigated the performance of full-blood Indian children and of white children on the California Achievement Test Batteries at three levels in order to study

> the degree of relationship between the subtests, to determine the number of independent factors which account for the scores on the different subtests, how heavily each of the subtests is weighted with the factors, and which factors appear as a whole for the different groups of subjects.[1]

The subjects were 5,502 full-blood Indian children and 9,386 white children in federal, public, and mission schools in six geographical areas of the United States. The age range of the Indian examinees was from 8 to 24 years and that of the white examinees was from 8 to 20. Pearson product-moment correlations were computed for the six variables on the basis of race, sex, and grade level. The resulting correlations were arranged into 12 correlation matrices which were factored by the centroid method. Three factors emerged on the first oblique rotation. A second rotation produced three new factors. On the second oblique rotation the third factor emerged as not sig-

---

[1]Omer J. Rupiper, "Multiple Factor Analysis of Academic Achievement: A Comparative Study of Full-blooded Indian and White Children," Journal of Experimental Education, 28 (1960), 177-205.

nificant for full-blood Indian male children. Factors A and C had characteristics of a "verbal nature" and Factor B, of a "numerical nature," with highly significant intercorrelations among the factors. Tentative conclusions suggested by this investigation were: two factors were found for full-blood Indian male children while three factors were found for each of the other nine groups, factor loadings for full-blood Indian male children were higher than for any other group, the verbal factor had a higher loading for full-blood Indian male children than for any other group, the factor loadings of a numerical nature were generally higher for full-blood Indian children than for the other groups, the factor patterns for all groups were more similar than otherwise, and a common second-order factor was evident in all groups.

The dimensions of the "meaning vocabularies" of children were investigated in 1954 by Russell.[1] This study sought to discover the breadth, depth, and height (grade level) of the vocabularies of children in the fourth through the twelfth grades. Nine tests were administered: seven specially designed vocabulary tests named Social Studies I, Social Studies II,

---

[1]David H. Russell, The Dimensions of Children's Meaning Vocabularies in Grades Four through Twelve ("University of California Publications in Education," No. 11; Berkeley: University of California Press, 1954), pp. 315-414.

Science I, Science II, Mathematics, Miscellaneous I, Miscellaneous II, and the SRA Primary Mental Abilities Test and the California Reading Test. The intercorrelations of the 14 variables were computed for six groups determined by sex and grade level (grades 4, 5, 6; grades 7, 8, 9; grades 10, 11, 12). The resulting coefficients were arranged into correlation matrices and subjected to factor analysis by Thurstone's centroid method. The projected factor loadings were arranged into factor matrices and rotated to simple structure.

The factor loadings extracted by rotation were approximately the same for grades 4, 5, 6, 7, 8, and 9. Analysis of these loadings suggested labeling Factor I as verbal meaning, Factor II as general mental ability, and Factor III as a "Reading to Follow Directions" factor. The verbal meaning factor had smaller projections for grades 10, 11, and 12 than for the lower grades but the other two factors were defined as approximately the same.

The general conclusions made from this investigation were: 1) the seven vocabulary tests were highly saturated with a verbal meaning factor, 2) there was a consistent improvement in vocabulary scores for both boys and girls from the fourth through the twelfth grade, 3) this growth was slower in grades 10, 11, and 12, 4) boys obtained higher scores than girls at

similar grade levels, 5) the attempt to measure breadth and depth of vocabulary yielded no significant inferences, and 6) the girls' vocabulary achievement became somewhat more specialized earlier than the boys'.

With respect to the California Reading Test, the variables yielded by performance on it projected significant loadings on the factors identified as verbal meaning, general mental ability, and reading to follow directions. At all levels investigated, the projections of this test were significantly higher on the verbal factor than on the other two factors. Moreover, this significance became progressively higher with the higher grade levels.

In 1955 Caffrey and Smith attempted to determine the number of factors which would account for the intercorrelations among the Davis-Eells Games, California Test of Mental Maturity, Occupational Ratings, and an experimental test of auding ability; and to identify the factors tentatively and assess their explanatory "value" in the analysis of so-called culture-fair tests of ability.[1] The subjects were 302 fifth grade pupils in six districts of Los Angeles County. The

---

[1]John Caffrey and Thomas W. Smith, "Preliminary Identification of Some Factors in the Davis-Eells Games," Paper read before the Annual Convention of the American Psychological Association, San Francisco, September 6, 1955.

tests used were: Davis-Eells Games, California Test of Mental Maturity, California Achievement Test, Auding Test, and Warner-Smith Occupational Ratings. The subtests of these instruments yielded 15 variables which were transformed into normalized stanine scores for the purpose of computing coefficients of correlation. The resulting correlation matrix was subjected to factor analysis by the centroid method. The factor loadings were arranged into a factor matrix which was rotated orthogonally by the Varimax method.

Three major factors appeared to account for the intercorrelations among the variables. Factor I was identified as the Davis-Eells "Problem Solving Index" factor since the Games subtests had the largest loadings on it. Since the subtests of the California Test of Mental Maturity and the California Achievement Test projected the largest loadings on Factor II, the latter was termed the Visual-Verbal Achievement factor. The third factor was identified as Auditory-Verbal since it appeared to represent the comprehension of spoken language. The Davis-Eells subtests contributed more to the loadings on Factor III than on Factor II. An oblique solution of the three factors produced a correlation of 0.67 between Factors I and III, Problem Solving and Auditory-Verbal, respectively.

These findings led to the following conclusions: 1)

problem-solving ability was a variable which involved auding ability and a visual perception factor, both of which may be sources of the "cultural bias" which persists in the Davis-Eells Games, 2) the Davis-Eells may be affected by factors other than reading and auding, and 3) socio-economic bias in the latter skills affects scores in a manner not consistent with the "culture-fair" criterion proposed by Davis-Eells.

As has been pointed out, the California Achievement Test and the California Test of Mental Maturity projected significantly high loadings only on Factor II (visual-verbal achievement). It would appear that the Factor should be identified as verbal-mathematical-spatial. Apparently, these two tests were measuring substantially the same patterns of abilities in this investigation. They might be identified as measures of scholastic ability or simply as measures of schooling.

In 1961 Wolins, MacKinney, and Stephans attempted to isolate and identify the factors contributing to performance on achievement in high school biology, chemistry, and physics; and on standardized measures of aptitude, interest, and achievement.[1] The subjects were students at Riverside-Brookfield High

---

[1]Leroy Wolins, A. C. MacKinney, and Paul Stephans, "Factor Analyses of High School Science Achievement Measures," Journal of Educational Research, 54 (1961), 173-77.

School, Riverside, Illinois, who had completed, at least, one semester of one or more of the three sciences under consideration. Three analyses were made: one for biology (N = 119), one for chemistry (N = 75), and one for physics (N = 87). The variables were derived from the raw scores on the Cooperative General Achievement Tests, Revised Series, Test II; California Achievement Test, Advanced Battery; California Short-Form Test of Mental Maturity, Verbal; California Short-Form Test of Mental Maturity, Advanced, Verbal; Cooperative English Test, Lower and Higher Levels, Test A, Mechanics of English; Cooperative American History Test; Nelson Biology Test; Anderson Chemistry Test; and Kuder Preference Record (Vocational), Scientific Scale; and grade point averages in biology for the first and second semester, and in physics for the first semester only. Sex constituted the sixteenth variable. All 16 variables, however, were not involved in each analysis. The coefficients of correlation were computed according to the product-moment method. The three correlation matrices were arranged and subjected to factor analysis by Thurstone's centroid method. The resulting factor loadings were arranged into factor matrices and rotated orthogonally for meaningfulness. The extraction of three factors for the biology and chemistry samples and of two factors for the physics sample accounted for all major relation-

ships among the variables since all residuals were below .11.

The first factor for each sample was identified as general intelligence because of the loadings from all the variables except sex and Kuder. The highest loadings were from the California Achievement Test. The second factor for each sample was identified by loadings from all of the science achievement tests, the physics grade, and very high loadings of sex and the Kuder Scientific Scale as a male interest-achievement factor. Only small loadings of the Cooperative American History Test and the Cooperative English Test were projected on this factor. The third factor was found only in the biology and chemistry samples. Small loadings of the English and History tests were found here, also. The investigators identified factor three as a specific achievement factor. The analysis of the saturations projected on the factors by the variables led to the conclusions that the Kuder Scientific Scale should probably not be used alone as a counseling tool since it appeared more a measure of masculinity-femininity than of scientific interest, that the invalid variance contained in science grades was a result of teachers including a discipline factor in the grades, and that standardized measures of achievement should be used instead of teachers' grades in analytical studies to eliminate the invalid variance.

It has been indicated that the California Achievement Test projected the highest loadings on factor one. Attention is called to the fact that the variables from the California Mental Maturity Test had significant projections only on factor one. Moreover, the only significant saturations for grade point average appeared on factors one and three, although even these were of minimal significance.

In 1950 a group of graduate students at the Catholic University of America, Washington, D.C., attempted to isolate and identify the factor patterns projected by performance on measurements of intelligence and achievement. The several members of the group planned to achieve this goal by specifically distinct approaches, each carrying out his own investigation based on the raw data derived from the administration of a battery of three intelligence tests and one achievement test to the same group of high school students in north-eastern Iowa. The subjects were 205 boys and 282 girls from the junior grade of 22 Catholic coeducational high schools. The tests used were: California Test of Mental Maturity, Non-Language, Advanced Series, Grade Nine-Adult; California Test of Mental Maturity, Language Section, Advanced Series, Age Eleven-Adult; Pintner General Ability Test, Verbal Series, Advanced Test, Form A; McManama Exercises in Cognitive Ability, Form A; and

Iowa High School Content Examination, Form L. The data yielded by the performance of the subjects on these tests supplied material for the several independent investigations.

Ruszel sought to determine whether the organization of mental abilities was the same for boys and girls or different.[1] Omitting the four variables derived from performance on the Iowa Test, Ruszel computed the coefficients of correlation of the remaining 26 variables by the product-moment method. Two correlation matrices, one for girls and one for boys, were constructed with the correlations. Each matrix consisted of 650 coefficients, all of which were positive or zero-order. The range for the girls was from 0.72 to 0.04 and for the boys, from 0.68 to 0.00. The two factor matrices were subjected to factor analysis by the centroid method. Each matrix yielded three factors which were then rotated both orthogonally and obliquely. The first factor extracted from each matrix was a numerical and spatial factor. This factor was the same for girls and boys. The second factor extracted from each matrix was a verbal and cognitive factor. This factor, also, was the same for both groups. The third factor to emerge was highly correlated with the verbal and cognitive fac-

---

[1]Humphrey Ruszel, Test Patterns in Intelligence (Washington, D.C.: The Catholic University of America Press, 1952).

tor. There were, however, some significant differences between the groups in performance on the subtests. Ruszel concluded from these findings that the factor patterns did not show a sex difference in mental abilities but that the differences in the performance of the two groups were the result of different combinations of abilities utilized in solving tasks. An additional observation which could be made on the basis of this study is that the projections of the California Test of Mental Maturity demonstrated that this test is a significantly meaningful measure of those abilities generally considered as intellective, viz., numerical, cognitive, verbal, and spatial.

Driscoll investigated the factor structure of mental abilities in boys.[1] Utilizing the data derived from performance on all the tests mentioned above, he computed the coefficients of correlation for the 30 variables yielded by the boys' performance on the subtests by the product-moment method. The resulting 870 coefficients of correlation were arranged into a correlation matrix with a range of 0.669 to 0.003. All the coefficients were positive or zero-order. The correlation ma-

---

[1]Justin A. Driscoll, Factors in Intelligence and Achievement (Washington, D.C.: The Catholic University of America Press, 1952).

trix was subjected to factor analysis by the centroid method. An extended factor matrix was determined from the resulting factor matrix. The latter suggested the construction of a transformation matrix which, in turn, resulted in a new factor matrix consisting of three factors: numerical, verbal, and cognitive. The numerical factor had a correlation of 0.306 with the verbal factor and of 0.505 with the cognitive factor. The verbal factor had a correlation of 0.427 with the cognitive factor. These findings were corroborated by subjecting the data to a second analysis utilizing the Spearman-Moore technique. In the second analysis, however, a fourth factor, identified as spatial, emerged. The spatial factor represented a more refined distribution of the loadings of the numerical factor. This analysis demonstrated the presence of a super-general factor among the four group factors. The multiple correlation coefficient was 0.936. Each group factor proved to be made up of two factors--the super-general factor and one general factor. Driscoll concluded that the covariance of the Iowa subtests was explained by one general factor which was highly related to verbal ability. Since the achievement tests contributed proportionately more to the common variance and correlated as well or better with the four group factors than did the mental ability tests, he also in-

ferred that the latter were no better measures of mental ability than were the former. One could further remark that the significant loadings projected by the California Test of Mental Maturity indicate that this test is heavily weighted with factors identified as numerical, cognitive, verbal, and spatial.

Doyle investigated the pattern of common variance measured by the forementioned tests and the significance of the groupings in achievement in four broad academic areas.[1] He limited his investigation to the data derived from the performance of the girls on the tests. The 30 subtest scores were correlated by the product-moment method. The resulting correlation matrix of 870 positive and zero-order coefficients ranging from 0.729 to 0.026 was factor analyzed by the centroid method. The centroid factor matrix was rotated orthogonally and obliquely. Three positively correlated factors were extracted: mathematical, verbal, and cognitive. The coefficient of correlation between the mathematical factor and the verbal factor was 0.454; between the mathematical factor and the cognitive factor, 0.559; and between the verbal factor and the cognitive factor, 0.691. These correlations led to the con-

---

[1]Andrew M. Doyle, Some Aspects of Ability and Achievement in High School Girls (Washington, D.C.: The Catholic University of America Press, 1952).

clusion that all the factors appeared to have the common characteristic of the presence of "symbols to represent meaning." The correlations also demonstrated that the differential achievement on the Iowa battery was dependent upon the verbal factor in all cases and that the mathematics subtest of the battery was also dependent on the mathematical factor. Attention is called to the fact that in this investigation, also, the California Test of Mental Maturity projected significant loadings on the mathematical, verbal, and cognitive factors.

Roesslein studied the comparative factor patterns of boys differentiated on the basis of achievement.[1] Classification as high and low was made in terms of the subjects' performance on the forementioned achievement tests. There were 80 subjects in each group. The variables derived from the 26 subtests of ability were correlated by the product-moment method for each group. Two correlation matrices of 650 positive and zero-order coefficients each were constructed. The range for the higher group was from 0.588 to -0.198 and for the lower group, from 0.629 to -0.265. These two correlation

---

[1]Charles G. Roesslein, Differential Patterns of Intelligence Traits Between High Achieving and Low Achieving High School Boys (Washington, D.C.: The Catholic University of America Press, 1953).

matrices were subjected to factor analysis by the centroid method. The unrotated factor matrices were rotated obliquely. Four factors were extracted from each matrix: cognitive, mathematical-spatial, verbal-spatial, and a second reasoning factor. Further analysis disclosed super-general factors the high intercorrelations of which seemed to indicate the presence of a third-order factor. Moreover, a higher degree of intercorrelation among the primary factors in the higher group was found. Roesslein attributed the latter fact to the possession of a greater amount of cognitive power by the higher group. In this investigation, too, the California Test of Mental Maturity projected highly significant loadings on the cognitive, mathematical, verbal, and spatial factors.

McCormick studied the possibility of differentiation in factor pattern in the performance on tests of mental ability and achievement of two groups of girls differentiated on the basis of cognitive ability.[1] The subjects were selected on the basis of their performance on the McManama Exercises in Cognitive Ability. From the total group of 282 girls reported on above, the investigator established two groups of

---

[1]Sister William Pauline McCormick, Factors of Intelligence in High and Low Cognitive Ability Groups (Washington, D.C.: The Catholic University of America Press, 1954).

100 each in terms of the top and bottom hundreds ranked in a frequency distribution of obtained scores. The product-moment method was used to determine the coefficients of correlation of the raw scores of the 30 subtests for the two groups. Two correlation matrices of 870 positive or zero-order coefficients each were constructed. The range for the lower group was from 0.582 to -0.178 and for the higher group, from 0.651 to -0.170. The correlation matrices were subjected to factor analysis by the centroid method. The factor loadings of both matrices were plotted on coordinate paper and subjected to oblique rotation. Transformation matrices were then obtained, and from these the final factor matrix saturations were calculated. Four factors were extracted from each matrix: verbal, numerical, cognitive, and spatial. Because all the intercorrelations of the extracted factors were positive and high, they were factored by orthogonal rotation. Two second-order factors were obtained for each group as a result of this analysis. These were identified as verbal and mathematical-spatial for the upper group and as cognitive and mathematical-spatial for the lower group. The analysis of the 30 subtests indicated that the four factors involved in performance on them could not be differentiated in the two groups, although their respective contributions to the common variances

were often dissimilar. In this study the California Test of Mental Maturity projected significant loadings on all the factors, however, the only projection of significance on Factor A (verbal) was the subtest on vocabulary.

In 1951 Elias investigated the extent to which personality attributes in addition to the cognitive are factors which enter into intelligence and achievement test scores of children in the sixth grade.[1] The subjects were two undifferentiated groups of 144 white children each with no marked deviations in age, physical health, or reading ability. The tests used were the Stanford Achievement Test, Intermediate, Form E, the SRA Primary Mental Abilities Test, Intermediate, Form AM, the new California Test of Mental Maturity, 1957 S-Form, Elementary, the Revised Beta Examination, and the California Test of Personality, Elementary. Performance on the subtests of these instruments yielded 33 variables for each group. Coefficients of correlation for each group were computed and the two resulting matrices of correlations were factor analyzed by the centroid method. The resulting factor

---

[1]Jack Z. Elias, "Non-Intellective Factors in Certain Intelligence and Achievement Tests: An Analysis of Factors in Addition to the Cognitive Entering into the Intelligence and Achievement Scores of Children at the Sixth Grade Level" (unpublished Doctor's dissertation, New York University, 1951).

matrices were subjected to orthogonal rotation. Separate analyses were made for each group.

Factors tentatively identified for the first group were: general, verbal, spontaneity, attention, and two whose meaning was less clear, pseudo-adjustment and reactivity; and for the second group, they were: general, verbal, attention, spontaneity, and two less clear, designated as reactivity and social maladjustment. These factors were classified on the basis of test content and/or previous factorial research as: general and verbal - intellective; attention, spontaneity, social maladjustment, pseudo-adjustment, and reactivity - non-intellective.

Attention appeared to be the major non-intellective factor projected by the four Stanford Achievement Tests, the California, and the Beta tests. Spontaneity was projected significantly by the SRA tests of Word Fluency and Number. Both of these factors entered into all the tests of intelligence as major projections whereas the other non-intellective factors were projected by the same tests with varying degrees of significance. The analysis of the variances of all the variables indicated that each of them measured the intellective and non-intellective factors in various degrees.

The conclusions made by Elias from this study were:

1) the Stanford Achievement Tests appeared to be better measures of the general factor than were the intelligence tests, 2) significant sex differences were found for some tests, and 3) there was need for further investigation of this nature. One could further observe that the California Test of Mental Maturity contributed significantly to the factor identified as intellective whereas the California Test of Personality made a similar contribution to the factor identified as non-intellective.

In 1963 Haller and Miller investigated the measureability of the level of occupational aspiration (LOA) of high school students in terms of their performance on the Occupational Aspiration Scale (OAS), designed to measure this factor.[1] Their study sought to develop a reliable, valid, and simple instrument for measuring LOA. The subjects were a randomized sample of 442 seventeen year old male students enrolled in the school system of Lenawee County, Michigan, in the spring of 1957. The instruments used in the study were: The Occupational Aspiration Scale, The 16 Personality Factor

---

[1]A. O. Haller and Irwin W. Miller, The Occupational Aspiration Scale: Theory, Structure and Correlates ("Department of Sociology and Anthropology Technical Bulletin 288"; East Lansing, Michigan: Michigan State University Cooperative Extension Service, 1963).

Test, Form B, Test of G-Culture Free-Scale 3 A, California Test of Personality, The MSU Work Beliefs Check-List, and a questionnaire, entitled Occupational Plans of Michigan Youth. Together with the grade point average for 1956-1957, these instruments yielded 36 variables, two of which were excluded from the investigation as irrelevant.

The coefficients of correlation of the 34 variables were computed by the product-moment method. The resulting correlation matrix was factor analyzed by the principal axes method. This process yielded only right principal axes of which three accounted for a substantial percentage of the total variance. These data were arranged into a factor matrix which was rotated to approximate orthogonal simple structure by the quartimax method. Factor I had high or moderate loadings on all the subtests of OAS and accounted for a high percent of the variance. The other two factors did not exhibit any systematic pattern nor did they account for any significant proportion of the variance. The test designers, therefore, concluded that the OAS was substantially a one factor test and that it was a reliable, stable, and approximately valid instrument.

In this study variable 20 was the total adjustment score of the California Test of Personality. A correlation

of 0.30 was found between variable 20 and LOA. This correlation and the other correlations of this test with the remaining variables as well as the factor loadings contributed by this test indicated an association of positive values of character, personality, and adjustment with the total adjustment score.

The findings and conclusions of the foregoing 12 selected factorial analyses are of importance to the present study since, as has been indicated, each of them used one or more of the standardized tests employed in this study. Although cross-references are made to these investigations in Chapter Four when the present factorial analysis is presented, the findings and conclusions are briefly summarized here.

Elias found seven factors which he classified on the basis of test content and examination of other factorial analyses as intellective and non-intellective.[1] Using this distinction, the factors isolated by the 12 forementioned studies may be considered as: intellective - general scholastic, general mental ability, general intelligence, cognitive, verbal, visual-verbal, auditory-verbal, verbal-spatial, mathematical, mathematical-spatial, numerical, numerical-spatial, spatial,

---

[1]Elias, op. cit.

scientific, literary, mechanical, problem-solving, male interest-achievement, and reading to follow directions; and non-intellective - personality, social maladjustment, pseudo-adjustment, "social butterfly," level of aspiration, attention, spontaneity, and reactivity.

The conclusions of these investigators were:

1. There is evidence of interrelationships between factors contributed to by measures of temperament and achievement.

2. Verbal and numerical factors are highly intercorrelated.

3. Vocabulary tests contribute highly to factors identified as verbal.

4. Problem-solving ability involves auding ability and visual perception.

5. Socio-economic bias appears to be present in reading and auding ability.

6. The Kuder Scientific Scale is more a measure of masculinity-femininity than of scientific interest.

7. Course grades contribute invalid variance as a result of teacher bias.

8. Standardized measures should be used in factorial analyses instead of teachers' grades to avoid invalid variance.

9. There appear to be no sex differences in mental abilities.

10. There appear to be no race differences between white children and full-blood Indian children in mental abilities.

11. Achievement tests appear to be as good as ability tests for measuring mental ability.

12. Intellective (cognitive) factors appear to have the common characteristic of the presence of "symbols to represent meaning."

13. Intellective (cognitive) factors appear to correlate higher in groups of boys of higher ability than in groups of boys of lower ability. There appears to be no differentiation of this kind in higher and lower groups of girls.

CHAPTER III

DESIGN OF THE STUDY

The meaningful interpretation of the findings of this investigation necessitates as thorough an understanding of the sources of the data analyzed as possible. This chapter, therefore, is dedicated to a discussion of the constituents of the sample and of the tests by which they were measured. For the purpose of further clarification a discussion of the methods by which the analysis of these data was made is also included here.

Subjects of the Sample

Sixty-five male students were selected as subjects for this factorial investigation upon entering the ninth grade at St. Gregory's High School, a college preparatory school for boys at Shawnee, Oklahoma. In September of the school year 1956-1957, two batteries of achievement tests of which one included a mental ability test and a religion test, a separate test of mental ability, a personality test, and a separate

religion test were administered to the sample. All were standardized tests and commercially available. The scores of 59 of the examinees were selected for use in the investigation. The six examinees excluded were eliminated on the assumption that their inclusion would contribute an undeterminable quantity of invalid variance to the results since they were not native speakers of the English language.

The examinees included in the sample constituted a substantially homogeneous group. All were entering ninth-grade male students. The age range of the group spread from thirteen years and three months to fifteen years and five months. Fifty students (85%) were between thirteen years and nine months and fourteen years and nine months of age. One student (2%) was below the latter range and eight students (13%) were above it. Fifty-five students (93%) had attended Catholic elementary schools for eight years. The remaining four students (7%) had attended Catholic elementary schools for at least four years.

The socio-economic background of the examinees was reasonably comparable. All came from families in which the "bread-winners" were classified as professional, proprietary, or managerial. All were from states in the southwestern section of the country. Visits to the homes of the examinees

indicated that all were living in circumstances typical of the upper-middle class of this geographical region. Thirty-four students (58%) were from large cities; eight students (13%) were from small cities; and 17 students (29%) were from rural areas. All were members of the Catholic Church. Fifty-seven students (97%) had been Catholics since infancy. Two students (3%) became members of the Catholic Church at the age of ten years. Thirty-six students (61%) were born of parents both of whom had always been Catholics. Nineteen students (32%) were born of parents one of whom joined the Catholic Church as an adult. Three students (5%) were born of parents both of whom became members of the Catholic Church as adults. One student (2%) was from a family in which one parent was a member of a Protestant denomination.

Tests Used in the Study

The tests employed in the study were selected for several reasons. All of them were standardized instruments. Separate answer sheets were available for them all. All could be scored by machine and by hand. The time limits for administration were not excessive for the age of the subjects used in this study. The instructions for administering the tests and the directions for taking the tests were of an appropriate level for this specific age group. The supervisor's manuals

contained adequate instructions for scoring and for interpreting the scores. The High School Placement Test was selected because it was the only battery with subtests of mental ability and basic skills which included a religion subtest. The Religion Test for High Schools was selected because it was a standardized measure of achievement in religious instruction. The California Achievement Tests and the California Test of Mental Maturity were selected as criteria because their subtests and the scores yielded by them were similar to the subtests of and the scores yielded by the High School Placement Test. The California Test of Personality was selected to determine whether any personal and/or social adjustment factors were related to performance in basic skills and religion as measured by the achievement batteries and the mental ability tests. A list of these five tests, consisting of 40 subtests, is given in Table 1 after which is found a description of each test as summarized from the data given in the respective manuals.

TABLE 1

TESTS USED IN THE STUDY

1. High School Placement Test, Battery for Entering Freshmen (Form A)
   - A. Area I: Mental Ability (110)[a]
     - Test 1. Mental Ability: Picture Similarities
     - 2. Mental Ability: Number Series
     - 3. Mental Ability: Inference

TABLE 1-Continued

4. Mental Ability: Arithmetic Reasoning
5. Mental Ability: Vocabulary

B. Area II: Reading (55)[a]
Test 6. Reading: Vocabulary
7. Reading: Recall of Information
8. Reading: Reading for Meaning

C. Area III: English (95)[a]
Test 9. English: Mechanics (Punctuation)
10. English: Mechanics (Capitalization)
11. English: Correct Usage

D. Area IV: Spelling (35)[a]
Test 12. Spelling

E. Area V: Arithmetic (41)[a]
Test 13. Arithmetic Reasoning

F. Area VI: Religion (55)[a]
Test 14. Religion

2. California Achievement Tests Complete Battery, Advanced, Grades 9 to 14, Form AA

A. Reading
Test 1. Reading Vocabulary (90)[a]
2. Reading Comprehension (55)[a]

B. Mathematics
Test 3. Mathematics Reasoning (60)[a]
4. Mathematics Fundamentals (80)[a]

C. Language
Test 5. Mechanics of English, and Grammar (80)[a]
6. Spelling (30)[a]

3. California Short-Form Test of Mental Maturity, Secondary, Grades 9 to 13, 1957 S-Form, Form A

A. Spatial Relationships Factor
Test 1. Sensing Right and Left (20)[a]
2. Manipulation of Areas (15)[a]

B. Logical Reasoning Factor
Test 3. Similarities (15)[a]
4. Inference (15)[a]

C. Numerical Reasoning Factor
Test 5. Number Series (15)[a]
6. Numerical Quantity (15)[a]

D. Verbal Concepts Factor
Test 7. Verbal Concepts (50)[a]

TABLE 1-Continued

4. California Test of Personality, 1953 Revision, Secondary, Grades 9 to College, Form AA
   - A. Personal Adjustment (90)[a]
     - Test 1A. Self-Reliance
     - 1B. Sense of Personal Worth
     - 1C. Sense of Personal Freedom
     - 1D. Feeling of Belonging
     - 1E. Withdrawing Tendencies
     - 1F. Nervous Symptoms
   - B. Social Adjustment (90)[a]
     - Test 2A. Social Standards
     - 2B. Social Skills
     - 2C. Anti-Social Tendencies
     - 2D. Family Relations
     - 2E. School Relations
     - 2F. Community Relations

5. Religion Test for High Schools, Form B
   - A. Religion Test (100)[a]

---

[a]Number represents highest possible score. Except in the case of the California Test of Personality for which there are no right or wrong responses, the score equals the number right.

The following summaries of data relevant to each test employed in the investigation were taken from the respective test manuals. All pertinent data are reported insofar as these are found in the manuals. For the sake of brevity, the names of the authors are omitted and the short titles of the tests are used.

High School Placement Test

The High School Placement Test comprises six major areas

measuring traits ordinarily associated with mental ability and achievement in reading, English, spelling, arithmetic, and religious instruction, respectively. The battery contains 391 items. The total actual testing time required is one hour and fifty-four minutes. The percentile norms are based on the performance of 2,132 eighth grade students representing 103 Catholic elementary schools and applying for admission to 17 different Catholic high schools.

The validity of the battery is reported in terms of content validity, construct validity, concurrent validity, and predictive validity. The mental ability test is reported to have a correlation of 0.77 with the Otis Beta Test and of 0.82 with the Terman-McNemar Test. In the second study the mental ability test produced a mean Intelligence Quotient of 104.62 with a sigma of 11.40 while the Terman-McNemar produced a mean Intelligence Quotient of 103.22 with a sigma of 9.35. Evidence of the predictive validity is presented in terms of correlations between the achievement tests of the battery and grades obtained at the end of the ninth grade in a study involving 163 cases. The English area was found to have a correlation of 0.53 with English grades; the spelling area, an r of 0.61 with spelling grades; and the arithmetic area, an r of 0.57 with arithmetic grades. Similar evidence

is given pertinent to the predictive validity of the mental ability test.

The reliabilities and related data are reported in Table 2.[1] Except in the case of the religion test, all re-

TABLE 2

RELIABILITY AND RELATED DATA FOR HIGH SCHOOL PLACEMENT BATTERY

Test Section	Reliability Estimates		S.D.	S.E. of Measurement
	Split-Half	Alternate Form		
Mental Ability	.93	.88	11.30[a]	3.91[a]
Reading				
Vocabulary	.79	.77	3.29	1.45
Rate	.78	.77	24.30	10.71
Comprehension	.83	.85	4.78	1.96
English	.84	.80	9.14	4.11
Spelling	.92	.90	8.62	2.76
Arithmetic	.80	.77	4.77	2.10
Religion	.75	b	8.40	4.20

[a]S.D. and S.E. for Mental Ability are in terms of I.Q.'s; all other data are in terms of raw scores.

[b]No alternate form of the religion test was in publication at the time the manual was published.

liabilities were determined by the split-half technique and the alternate form technique. The reliability of the religion

[1]O. F. Anderhalter, R. Colestock, and R. S. Gawkowski, "High School Placement Test, Battery for Entering Freshmen (Form A)." Scholastic Tests Manual (Chicago: Scholastic Testing Service Inc., 1955), p. 16.

test was found by means of the split-half technique.

California Achievement Tests Complete Battery

The California Achievement Tests Complete Battery was constructed to measure student achievement in fundamental reading, mathematics, and language skills and to provide a basis for planning remedial instruction in the areas in which individual students proved to be deficient. The battery yields scores in six areas: reading vocabulary, reading comprehension, mathematics reasoning, mathematics fundamentals, mechanics of English and grammar, and spelling. There are 395 items in the battery. The total actual testing time is two hours and forty-one minutes. The normative population was a randomized sample of over 100,000 cases.

The Manual contains very little data on the validity of the battery.[1] There is, however, a statement which ascribes content and construct validity to the battery. Test users are referred to the 1957 Technical Report on the California Achievement Tests for the necessary statistical data.[2]

---

[1]E. W. Tiegs and W. W. Clark, "California Achievement Tests Complete Battery, Advanced, Grades 9 to 14, Form AA," California Test Bureau Manual (Los Angeles: California Test Bureau, 1951).

[2]1957 Technical Report on the California Achievement Tests (Monterey, California: California Test Bureau, 1957).

The latter publication dedicates several pages to a display of validity data most of which are reported in tables of product-moment validity correlations between the six areas of the battery and the several levels of the California Test of Mental Maturity. Correlations between the language part of the CTMM and the area tests are reported as: reading vocabulary 0.81, reading comprehension 0.77, mathematics reasoning 0.64, mathematics fundamentals 0.57, mechanics of English 0.63, and spelling 0.47. Correlations between the non-language part of the CTMM and the area tests are reported as: reading vocabulary 0.59, reading comprehension 0.63, mathematics reasoning 0.63, mathematics fundamentals 0.58, mechanics of English 0.47, and spelling 0.29.

All the reliabilities for the battery were computed by the alternate form method. The reliabilities are reported in Table 3.[1]

California Short-Form Test of Mental Maturity

The California Short-Form Test of Mental Maturity is an instrument designed to measure mental development or capacity. The test comprises seven subtests devised to yield data on four factors: spatial relationships, logical reason-

---

[1]Tiegs and Clark, California Test Bureau Manual, p. 5.

TABLE 3

RELIABILITY AND RELATED DATA FOR CALIFORNIA ACHIEVEMENT TESTS COMPLETE BATTERY

Test Area	Alternate Form Reliabilities	S.E. of Measurement
Reading Vocabulary	.93	0.59
Reading Comprehension	.90	0.66
Total Reading	.92	0.62
Mathematics Reasoning	.91	0.73
Mathematics Fundamentals	.93	0.68
Total Mathematics	.93	0.66
Mechanics of English	.89	0.74
Spelling	.84	0.88
Total Language	.92	0.63
Total Battery	.98	0.33

ing, numerical reasoning, and verbal concepts. There are 145 items in the test. The test yields scores on language mental age, non-language mental age, and total mental age. Fifty-two minutes are required for the administration of the test. The test was standardized in a two-stage program on a randomized population of 65,000 subjects.

The Manual contains data on the validity of this test in terms of concurrent validity, construct validity, and predictive validity.[1] Concurrent validity is demonstrated by

[1]W. W. Clark and E. W. Tiegs, "California Short-Form Test of Mental Maturity, Grades 9 to 11, Secondary, 1957 S-Form, Form A," California Test Bureau Manual (Los Angeles: California Test Bureau, 1957), pp. 5-8.

consistently high correlations between this test and the Stanford-Binet which have been found to range from 0.81 to 0.88 in several studies. Additional evidence of concurrent validity is the correlation reportedly found between this test and the Wechsler-Bellevue. This is reported to be 0.81. The 1957 Technical Report on the California Test of Mental Maturity reports still further data on concurrent validity.[1] The Short-Form language section is reported to have a correlation of 0.79 with the Otis-Gamma Test and 0.99 with the Terman-McNemar test, whereas the non-language section is reported to have correlations of 0.95 with the former and 0.82 with the latter. Construct validity is demonstrated by several tables which clearly show the internal consistency of the several parts of the test. Neither the Manual nor the Technical Report contains satisfactory data on predictive validity. The authors simply report that the test has been used for years "in 'readiness' and other predictive studies."[2] They further observe that the test compares "very favorably with the Stanford-Binet and the two

---

[1] 1957 Technical Report on the California Test of Mental Maturity (Los Angeles: California Test Bureau, 1957), p. 29.

[2] Clark and Tiegs, California Test Bureau Manual, p. 7.

Wechsler tests for predicting successful performance in school, on the job, and other criteria where intelligence is an important variable."[1]

The reliabilities of the test were computed by the Rulon split-half technique. The data, based on the performance of 200 examinees selected from the standardization sample, are reported in Table 4.[2]

TABLE 4

RELIABILITY AND RELATED DATA FOR CALIFORNIA SHORT-FORM TEST OF MENTAL MATURITY

Data	Reliability Coefficient	Mental Age		Intelligence Quotient	
		S.E. of Measurement	S.D.	S.E. of Measurement	S.D.
Language	.94	9.5	38.9	3.9	16.0
Non-Language	.92	7.8	27.5	4.5	16.0
Total	.95	6.5	29.1	3.6	16.0

California Test of Personality

The California Test of Personality, Secondary, Grades 9 to College, Form AA, comprises two parts, personal and social adjustment, designed "to identify and reveal the status

[1] Ibid.

[2] Ibid., p. 4.

of certain highly important factors in personal and social adjustment"[1] and "to provide the data for aiding individuals to maintain or develop a normal balance between personal and social adjustment."[2] The first part is a measure of personal security. It consists of six component parts each containing 15 items: self-reliance, sense of personal worth, sense of personal freedom, feeling of belonging, withdrawing tendencies, and nervous symptoms. The second part is a measure of social security and consists, also of six component parts each with 15 items: social standards, social skills, anti-social tendencies, family relations, school relations, and community relations. The total number of items is 180. There is no limit for the administration of the inventory although the Manual suggests 45 minutes with the caution that each examinee should be permitted to finish. The norms for this form of the instrument were derived from data secured on 3,331 students in grades 9 to 14 inclusive in schools in Connecticut, Massachusetts, Michigan, Pennsylvania, South Dakota, and California. Measures were taken to secure a sta-

---

[1]L. P. Thorpe, W. W. Clark, and E. W. Tiegs, "California Test of Personality, 1953 Revision, Secondary, Grades 9 to College, Form AA," California Test Bureau Manual (Los Angeles: California Test Bureau, 1953), p. 2.

[2]Ibid.

tistically randomized sample.

The discussion of the validity of the instrument in the Manual contains no correlation data. Content validity is demonstrated by a description of the procedure involved in the selection of items by a team of judges from monographs written by psychologists and clinical psychologists. Predictive validity is ascribed on the strength of statements made by teachers, reports made by the New York City Schools, and research at Syracuse University which found the test correlated more closely with clinical findings than any other personality test. Construct validity is based upon one research report involving 168 children from the sixth grade. According to this report "the test revealed a statistically significant difference in favor of the better-adjusted group . . ."[1] Also, the Manual contains the statement that "some 90 [studies] testify to the value of this instrument . . ." and that "many clinical psychologists consider the California Test of Personality a very useful instrument in obtaining controlled interview data . . ."[2]

The reliability coefficients of the inventory were com-

---

[1] Ibid., p. 8.

[2] Ibid.

puted by the Kuder-Richardson formula. The Manual contains tables of these coefficients for each of five levels of the test. In the case of the form under consideration, the reliability coefficients were based on data from a population of 2,262 subjects. The personal adjustment part of the test has a reliability r of 0.90 with a S.E. of Measurement of 3.72 and the social adjustment part has a reliability r of 0.89 with a S.E. of Measurement of 3.48. The reliability r for the total inventory is reported as 0.93 with a S.E. of Measurement of 5.56.

Religion Test for High Schools

The Religion Test for High Schools was designed "to measure the high-school student's knowledge of those fundamentals in moral and dogmatic theology which are so important that they should not under any circumstances be neglected in a high-school religion course."[1] The 100 multiple-choice items of the test are divided into three areas covering the main divisions of religious knowledge: the Creed, 35 items; the Commandments, 35 items; and the Means of Grace, 30 items. Forty-five minutes are alloted for the administration of the

---

[1]A. G. Schmidt and O. F. Anderhalter, "Religion Test for High Schools," General Manual (Chicago: Loyola University Press, 1955), p. 1.

test. Norms are based on a randomized sample of 6,526 subjects representing 29 high schools located in 10 different states.

The Manual contains a discussion of content and construct validity. No data relevant to predictive validity are presented since the test authors did not consider this germane to their objectives. It was not feasible to determine concurrent validity since the only other standardized test of religion available, the Religion Essentials Test by Austin G. Schmidt, S.J., was so different "in purpose, content, and structure that it did not seem safe to use one as a criterion of the validity of the other."[1] Content validity appears to be well established in that the items in the test were determined by a group of three hundred theologians, supervisors, and teachers of religion throughout the nation. Construct validity is demonstrated by data gathered on more than 5,000 subjects in an experimental testing and analysis of the items.

The reliabilities of the test were computed by utilizing the split-half technique and the alternate form technique. The median split-half reliability r for the total

---

[1] Ibid.

test was found to be 0.87 with a range of 0.83 to 0.88 for all forms of the test. The median alternate form reliability r for the total test was 0.81. The total test of the form included in this study was found to have a mean of 52.34 with a sigma of 15.30 and a S.E. of Measurement of 6.66.

Analytical Procedure

The raw scores on the five tests together with the first semester grade point average in ninth-grade religion constituted the data for this analysis. Although the five tests employed in this study had a total of 40 subtests, the subtest scores were not used as variables in all cases. This plan was dictated by two statistical principles pertinent to a study of factorial analysis. These principles are: 1) the number of variables to be submitted to factorial analysis should be, at least, less than one half the size of the number in the sample, and 2) the reliability of the subtotal score is higher than the reliability of the score for each subtest. Hence, the raw scores of the 14 subtests of the High School Placement Test, henceforth referred to as HSPT, were reduced to eight variables by employing only the total scores of the mental ability test (5 subtests) and of

the English test (3 subtests). The 12 subtest scores of the California Test of Personality, henceforth referred to as CTP, were reduced to two variables by using only the two subtotal scores (personal adjustment and social adjustment), each comprising six subtests. The seven subtests of the California Short-Form Test of Mental Maturity, henceforth referred to as CTMM, were reduced to two variables representing the language and non-language totals. The total score yielded by the Religion Test for High Schools, henceforth referred to as RTHS; the first semester grade point average in ninth-grade religion, henceforth referred to as GPA; and the six subtest scores yielded by the California Achievement Tests Complete Battery, henceforth referred to as CAT, constituted eight additional variables. This procedure resulted in the determination of 20 variables to be submitted to factorial analysis. The designation of these 20 variables by number and by name is reported in Table 5 together with the means, standard deviations, and standard errors of the means for the sample. Pearson product-moment correlations among the 20 variables were calculated and arranged into a correlation matrix. The correlation matrix was factor analyzed according to Thurstone's centroid method and rotated to simple structure by the Oblimax

TABLE 5

DESIGNATION BY NUMBER AND NAME OF THE VARIABLES IN THIS STUDY; MEANS, STANDARD DEVIATIONS, AND STANDARD ERRORS OF THE MEANS FOR THE SAMPLE

Test	Variables Number	Name	Mean	S.D.	S.E. of Means
1.		High School Placement Test (HSPT)			
	X_1	Mental Ability	72.14	14.41	1.89
	X_2	Reading: Vocabulary	15.15	4.31	.57
	X_3	Reading: Rate	88.22	2.58	.34
	X_4	Reading: Comprehension	25.07	6.44	.85
	X_5	Arithmetic	13.14	5.46	.72
	X_6	English	66.12	12.60	1.65
	X_7	Spelling	24.73	6.98	.92
	X_8	Religion	33.88	10.38	1.36
2.		California Test of Personality (CTP)			
	X_9	Personal Adjustment	65.41	12.40	1.63
	X_{10}	Social Adjustment	65.41	12.84	1.69
3.		California Short-Form Test of Mental Maturity (CTMM)			
	X_{11}	Language	26.59	8.70	1.14
	X_{12}	Non-Language	33.73	4.74	.61
4.		Religion Test for High Schools (RTHS)			
	X_{13}	Religion	39.37	14.19	1.86
5.		Grade Point Average (GPA)			
	X_{14}	GPA	2.41	0.91	.12
6.		California Achievement Tests Complete Battery (CAT)			
	X_{15}	Reading: Vocabulary	37.15	12.95	1.70
	X_{16}	Reading: Comprehension	29.42	7.84	1.03
	X_{17}	Mathematics: Reasoning	30.61	9.53	1.25
	X_{18}	Mathematics: Fundamentals	40.47	19.81	1.94
	X_{19}	Mechanics of English and Grammar	49.02	13.68	1.80
	X_{20}	Spelling	11.44	6.04	.79

method.[1] The variance of each loading on each factor, the communalities of the variances, the total variances, and the per cent of variance were computed for each factor.

The item analysis of the HSPT religion subtest was made according to the technique designed by Davis whereby the level of difficulty and the discriminating power of each item were computed.[2]

---

[1]L. L. Thurstone, Multiple Factor Analysis (Chicago: University of Chicago Press, 1931), pp. 149-75.

[2]F. B. Davis, Item-Analysis Data (Cambridge: Graduate School of Education, Harvard University, 1949).

CHAPTER IV

STATISTICAL ANALYSIS

Raw scores were obtained through the administration of the HSPT, the CTP, the CTMM, the RTHS, and the CAT to the sample comprised of 59 ninth-grade high-school boys as described in Chapter Three. While the sample was admittedly small, it could be assumed to be sufficiently devoid of bias to permit the making of meaningful inferences.

The data presented here are analyses of the relationships among the 20 variables in terms of zero-order correlations and the factor analysis of the resulting correlation matrix. As pointed out in Chapter Three, 19 variables were obtained from the performance of the sample on the five standardized tests mentioned above, and one variable consisted of the sample's first semester grade point average in ninth-grade religion. The data were processed by members of the staff of the Computer Laboratory at the University of Oklahoma.

Correlation Matrix

The obtained 190 Pearson product-moment coefficients

of correlation were arranged into a correlation matrix. These coefficients of correlation, ranging in size from r = 0.14 to r = 0.84, are presented in Table 6 along with the means and standard deviations for the respective variables.

The t test was used to determine the significance of the coefficients of correlation.

> In testing the null hypothesis for a coefficient of correlation, the required t is estimated by the formula
>
> $$t = r\sqrt{\frac{N - 2}{1 - r^2}}$$
>
> where r = obtained coefficient of correlation and N = number of pairs of observations.[1]

Thus, in this sample of 59, an r of 0.26 or greater proved to be significant at the five per cent level of significance and an r of 0.33 or greater proved to be highly significant at the one per cent level of significance.

$$t = 0.26\sqrt{\frac{59 - 2}{1 - (.26)^2}} = 2.03$$

and

$$t = 0.33\sqrt{\frac{59 - 2}{1 - (.33)^2}} = 2.63$$

---

[1]J. P. Guilford, Fundamental Statistics in Psychology and Education (New York: McGraw-Hill Book Co., 1956), p. 219.

TABLE 6

INTERCORRELATIONS OF RAW SCORES, MEANS, AND STANDARD DEVIATIONS OF VARIABLES FOR NINTH-GRADE STUDENTS (N = 59)[a]

Variables	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}
1 HSPT IQ		71	42	64	64	83	73	75	31	46	71	50	71	72	72	82	74	76	74	74
2 HSPT Reading Voc.	71		36	46	43	68	51	66	21	38	43	29	68	40	56	72	48	49	54	46
3 HSPT Reading Rate	42	36		59	32	31	49	34	19	20	49	20	51	24	62	54	33	33	40	48
4 HSPT Reading Comp.	64	46	59		44	53	57	55	22	26	63	42	57	50	57	66	47	56	48	53
5 HSPT Arithmetic	64	43	32	44		57	55	58	27	43	63	37	62	59	58	61	63	69	67	63
6 HSPT English	83	68	31	53	57		65	74	19	34	55	40	64	59	60	72	64	67	67	59
7 HSPT Spelling	73	51	49	57	55	65		73	15	30	65	47	58	56	72	69	54	64	76	75
8 HSPT Religion	75	66	35	55	58	74	73		14	38	54	47	74	59	56	69	68	69	70	62
9 CTP Personal Adjust.	31	24	19	22	27	19	15	14		75	16	51	21	30	22	18	21	32	26	24
10 CTP Social Adjust.	46	38	20	26	43	34	30	38	75		20	14	27	39	25	31	41	49	31	31
11 CTMM Language	71	43	49	63	63	55	65	54	16	20		37	73	55	80	73	59	64	67	75
12 CTMM Non-Language	50	29	20	42	37	40	47	47	51	14	37		45	53	37	42	44	40	50	44
13 RTHS Religion	71	68	51	57	62	64	68	74	21	27	73	45		61	80	80	66	69	78	73
14 Grade Point Average	72	40	24	50	59	59	56	59	30	39	55	53	61		55	55	71	77	73	57
15 CAT Reading Voc.	72	56	62	57	58	60	72	56	22	25	80	37	80	55		84	69	61	73	73
16 CAT Reading Comp.	82	72	54	66	61	72	69	69	18	31	73	42	80	55	84		64	65	67	67
17 CAT Math. Reasoning	74	48	33	47	63	64	54	68	21	41	59	44	66	71	69	64		76	71	61
18 CAT Math. Fund.	76	49	33	56	69	67	64	69	32	49	64	40	69	77	61	65	76		73	65
19 CAT Mech. of Eng.	74	54	40	48	67	67	76	70	26	31	67	50	78	73	73	67	71	73		79
20 CAT Spelling	74	46	48	53	63	59	75	62	24	31	75	44	73	57	73	67	61	65	79	
Mean	72.14	15.15	88.22	25.07	13.14	66.12	24.73	33.88	65.41	65.41	26.59	33.73	39.37	2.41	37.15	29.42	30.61	40.47	49.02	11.44
S.D.	14.41	4.31	2.58	6.44	5.46	12.60	6.98	10.38	12.40	12.84	8.70	4.64	14.19	0.91	12.95	7.84	9.53	14.81	13.68	6.04

[a]Decimals omitted.

In this study 18 coefficients of correlation, slightly less than 9½ per cent, were non-significant; 14, slightly less than 7½ per cent, were significant at the five per cent level of significance; and 158, slightly more than 83 per cent, were highly significant at the one per cent level of significance. Thus, over 90 per cent of the coefficients of correlation were accepted as significantly different from zero at or beyond the five per cent level of significance.

Sixteen of the 18 non-significant coefficients of correlation accounted for the relationships between the two variables yielded by the CTP and other variables. Eleven of the 14 coefficients of correlation significant only at the five per cent level of significance also accounted for the relationships between the two variables yielded by the CTP and other variables. Hence, apart from the CTP variables, there were only five coefficients of correlation which were not highly significant at the one per cent level of significance, and four of these were the coefficients of correlation between variable X_3 (HSPT reading rate) and four other variables: X_5 (HSPT arithmetic), X_6 (HSPT English), X_{14} (GPA), and X_{12} (CTMM non-language), the latter accounting for the fifth such correlation with an r of 0.29 with variable X_2 (HSPT reading vocabulary) significant only at the five per

cent level of significance.

The lowest coefficient occurred in two cases accounting for the relationships between variable X_8 (HSPT religion) and variable X_9 (CTP personal adjustment) and between variable X_{10} (CTP social adjustment) and variable X_{12} (CTMM nonlanguage). The coefficients of variables X_9 and X_{10} (CTP personal and social adjustment) were consistently low, as noted above, except in the relationship between themselves which was a highly significant r of 0.75. The highest coefficient occurred between variables X_{15} (CAT reading vocabulary) and X_{16} (CAT reading comprehension).

Concerning the implications of the obtained coefficients of correlation, Thurstone states,

> Certain features of the test configuration can be inferred by mere inspection of the correlation coefficients. If all the correlations are positive or zero, we can infer that all the test vectors are separated by acute angles. The largest angular separation would then be a right angle. It follows that the whole test configuration must be obtained within a cone of a 45 degree generating angle when the correlation matrix has no negative entries. This is the situation with psychological tests of the cognitive or intellective functions, which seem never to show any significant negative correlations.[1]

Examination of the correlation matrix revealed that all coefficients were positive. Therefore, the inferences that

---

[1]Thurstone, Multiple Factor Analysis, p. 93.

the configuration of the variable vectors did not spread beyond a right angle and that the variable vectors were contained within a 45 degree generating angle were made. These observations demonstrated that a common factor or factors were involved in performance on the variables and that, as observed by Thurstone above, the variables were probably weighted with a cognitive or intellective factor.

Centroid Factor Analysis

The correlation matrix was submitted to factor analysis by Thurstone's centroid method using a modified procedure (Oblimax method) adapted to IBM processing. The procedure of factorization posed the anticipated problem as to the number of factors to be extracted. Since Fruchter recommends Humphrey's Rule as best suited to determine the number of factors for samples that are not too large, this criterion was employed in the factorization procedures of this investigation.[1] The formula employed was, simply, that the product of the two highest loadings on a factor should exceed the standard error of zero r. The latter was computed by the formula

---

[1]Benjamin Fruchter, Introduction to Factor Analysis (New York: Van Nostrand Co., 1954), p. 79.

$$\sigma_{r_o} = \frac{1}{\sqrt{N - 1}}$$

in which N = size of sample. The standard error of zero r was found to be .131. The products of the two highest loadings on each factor were: K1 = .787, K2 = .240, K3 = .169, K4 = .130, and K5 = .054. Although the products obtained on factors K4 and K5 failed to meet the criterion for sufficient factors, the procedure of extraction was continued through these two factors.

> Since the criteria are not exact, it is frequently necessary, in practice, to extract more than the minimum number of factors indicated by the criteria and determine by rotation how many factors are needed to give a meaningful solution. If the rotations are started with too many factors, some of them may be 'residualized' (i.e., achieve small loadings only, usually between $\pm$ 0.20) during rotation.[1]

Moreover, the process of factorization was not carried beyond five factors since the residuals were significantly small. Also, the per cent of variance accounted for by the five factors extracted was found to be 99%. The remaining 1% of variance was assumed to be accounted for by error variance.

Unrotated Factor Matrix

The factor loadings projected by the process of fac-

---

[1] Ibid., p. 84.

torization were arranged into a factor matrix of unrotated factors as shown in Table 7. The variance of each loading on each factor was computed by obtaining the square of the factor saturation. The communalities (h^2) of the variances of the variables were determined by summing their respective variances on each factor, and the total variance of each factor was determined by summing the variance of each variable on the respective factors and combining these totals. The per cent of variance of each factor was computed by dividing the total variance into the individual factor variance. The results of these computations indicated that the first factor (K1) contributed 77% of the variance, the second factor (K2) contributed 8% of the variance, the third factor (K3) contributed 6% of the variance, the fourth factor (K4) contributed 5% of the variance, and the fifth factor (K5) contributed 3% of the variance. These five factors accounted for 99% of the total variance.

Several observations relevant to the unrotated factor matrix were considered necessary for the better elucidation of the meaningfulness of the rotated factor matrix to be considered later. Adcock states that "low reliability depresses the general level of correlation and so reduces the factor loading. Ratings of personality qualities usually have very

TABLE 7

UNROTATED FACTOR MATRIX[a]

Variables	Factors										h^2
	K1		K2		K3		K4		K5		
	Load.	Var.	Load.	Var.	Load.	Var.	Load.	Var.	Load.	Var	
1 HSPT IQ	912	832	-122	015	070	005	154	023	049	002	.877
2 HSPT Reading Voc.	691	477	-086	007	-211	045	388	151	147	022	.702
3 HSPT Reading Rate	541	293	378	143	-370	137	-055	003	-109	012	.588
4 HSPT Reading Comp.	698	487	194	038	-130	017	091	008	-124	015	.565
5 HSPT Arithmetic	741	550	-076	006	117	014	-178	032	099	010	.612
6 HSPT English	795	632	-159	025	117	014	312	100	111	012	.783
7 HSPT Spelling	806	650	213	045	135	018	066	004	-205	042	.759
8 HSPT Religion	806	650	-095	009	156	024	298	089	-068	005	.777
9 CTP Personal Adjust.	363	132	-401	161	-444	197	-337	114	-234	055	.659
10 CTP Social Adjust.	495	245	-600	360	-382	146	-244	060	-208	043	.854
11 CTMM Language	786	618	313	098	043	002	-140	020	165	027	.765
12 CTMM Non-Language	527	278	061	004	231	053	110	012	-225	050	.397
13 RTHS Religion	857	734	182	033	-035	001	046	002	109	012	.782
14 Grade Point Average	760	578	-171	030	258	067	-120	014	-062	004	.693
15 CAT Reading Voc.	836	699	334	112	-115	013	-128	016	228	052	.892
16 CAT Reading Comp.	864	746	164	027	-153	023	213	045	234	055	.896
17 CAT Math. Reasoning	793	629	-148	022	184	034	-121	015	171	030	.730
18 CAT Math. Fund.	835	697	-210	044	169	029	-139	019	058	003	.792
19 CAT Mech. of Eng.	854	729	053	003	238	057	-122	015	-049	002	.806
20 CAT Spelling	818	669	209	044	139	019	-157	025	-053	003	.760
Total Variance	(11.325)		(1.226)		(.915)		(.767)		(.456)		14.689
Per cent of Variance	77%		8%		6%		5%		3%		99%

[a]Decimals omitted.

low reliability . . ."[1] Attention was called above to the fact that the coefficients of correlation of variables X_9 and X_{10} (CTP personal and social adjustment) were consistently low. Attention is called here to the fact that these same two variables projected the lowest loadings on Factor one (K1) and that they are the only variables which appeared negatively on all the other factor loadings.

Thurstone remarks that "A zero factor loading means that the factor in question does not contribute to the variance of the test . . . that is, the factor is not involved in the test variance."[2] No zero loadings appeared on Factor one (K1). Five such loadings appeared on Factor two (K2). These were projections from variables X_2 (HSPT reading vocabulary), X_5 (HSPT arithmetic), X_8 (HSPT religion), X_{12} (CTMM non-language), and X_{19} (CAT mechanics of English and grammar). Three zero loadings appeared on Factor three (K3). These were projections of variables X_1 (HSPT mental ability), X_{11} (CTMM language), and X_{13} (RTHS religion). Four zero loadings appeared on Factor four (K4). These were projected by variables X_3 (HSPT reading rate), X_4 (HSPT read-

---

[1]C. J. Adcock, Factorial Analysis for Non-Mathematicians (London: Cambridge University Press, 1954), p. 77.

[2]Thurstone, Multiple Factor Analysis, p. 341.

ing comprehension), X_7 (HSPT spelling), and X_{13} (RTHS religion). Seven zero loadings appeared on Factor five (K5). These were projections of variables X_1 (HSPT mental ability), X_5 (HSPT arithmetic), X_8 (HSPT religion), X_{14} (grade point average), X_{18} (CAT mathematics fundamentals), X_{19} (CAT mechanics of English and grammar), and X_{20} (CAT spelling).

Fruchter points out that "Loadings of .2 or less are usually regarded as insignificant, loadings of .2 to .3 as low, .3 to .5 as moderate, .5 to .7 as high, and above .7 as very high."[1] Moreover, Thurstone states,

> The naming of a factor cannot be made with confidence unless the projections are as large as .50 and .60 so that the factor accounts for a fourth or a third of the variance of a test. Confidence in naming a factor is also determined by the number of tests that have significant projections of .40 or higher on the factor.[2]

Applying these criteria in the examination of the unrotated factor matrix, it was observed that on Factor one (K1) there were no non-significant loadings, whereas Factor two (K2) had only eight significant loadings: X_3 (HSPT reading rate), X_7 (HSPT spelling), X_9 (CTP personal adjustment), X_{10} (CTP social adjustment), X_{11} (CTMM language), X_{15} (CAT reading vo-

---

[1]Fruchter, op. cit., p. 151.

[2]L. L. Thurstone, Primary Mental Abilities ("Psychometric Monographs," No. 1; Chicago: University of Chicago Press, 1938), p. 12.

cabulary), X_{18} (CAT mathematics fundamentals), and X_{20} (CAT spelling); Factor three (K3), seven: X_2 (HSPT reading vocabulary), X_3 (HSPT reading rate), X_9 (CTP personal adjustment), X_{10} (CTP social adjustment), X_{12} (CTMM non-language), X_{14} (grade point average), and X_{19} (CAT mechanics of English and grammar); Factor four (K4), six: X_2 (HSPT reading vocabulary), X_6 (HSPT English), X_8 (HSPT religion), X_9 (CTP personal adjustment), X_{10} (CTP social adjustment), and X_{16} (CAT reading comprehension); and Factor five (K5), six: X_7 (HSPT spelling), X_9 (CTP personal adjustment), X_{10} (CTP social adjustment), X_{12} (CTMM non-language), X_{15} (CAT reading vocabulary), and X_{16} (CAT reading comprehension). Furthermore, according to Thurstone, only Factor one (K1) had loadings sufficiently large enough to permit the consideration of identifying its nature. This was not attempted, however, at this point, although some factorial analyses might attempt it, since, as Thurstone points out, "the configuration must be rotated before the factors can be even expected to have psychological meaning. One cannot expect to find several meaningful factors by any of the current methods of factoring a correlational matrix without rotation of the reference frame."[1]

---

[1] Ibid., p. 90.

Several additional observations pertinent to the unrotated factor matrix were made. All 20 variables were projected significantly on Factor one (K1). Variables X_9 (CTP personal adjustment) and X_{10} (CTP social adjustment) appeared significantly on all five factors, although the sign was consistently negative on all but Factor one (K1). Variables X_1 (HSPT mental ability), X_4 (HSPT reading comprehension), X_5 (HSPT arithmetic), X_{13} (RTHS religion), and X_{17} (CAT mathematics reasoning) appeared significantly only on Factor one (K1). All but Factor one (K1) were bipolar factors. According to Adcock, the foregoing observations indicated that the centroid factor matrix obtained in this analysis was typical and demonstrated the necessity of rotating the factor matrix to simple structure.[1]

Rotated Factor Matrix

Further consideration of the unrotated factor matrix was not productive of any additional insights as to the nature and structure of the abilities utilized in performance on the variables. Thurstone's principle of simple structure calls for the careful rotation of the factor axes to gain these insights. Attaining simple structure consists simply

---

[1]Adcock, op. cit., p. 30.

in rotating the axes of the factors so that the number of negative loadings is minimized and the number of zero loadings is maximized.

Oblique rotation was accomplished by means of the Oblimax method. The resulting factor loadings were arranged into a rotated factor matrix as presented in Table 8. Examination of the rotated factor matrix according to the same criteria considered above under the discussion of the unrotated factor matrix led to the following observations.

The loadings on Factor one (K1) were all small except for variable X_{14} (grade point average) which was highly significant with a saturation of .675. The abilities represented by this Factor appeared to have nothing in common with the basic skills tests and very little with the personality measures which had minimally significant projections of .135 for X_9 (CTP personal adjustment) and .149 for X_{10} (CTP social adjustment) on this Factor. A projection of .198 from variable X_3 (HSPT reading rate) indicated that the abilities represented by this Factor had a minimally significant relationship with this basic skill. On the basis of Thurstone's criterion for naming a factor, no attempt was made to identify Factor one. This Factor accounted for 7% of the total variance contributed to the relationships of the variables.

TABLE 8

ROTATED FACTOR MATRIX[a]

Variables	Factors										h^2
	K1		K2		K3		K4		K5		
	Load.	Var.	Load.	Var.	Load.	Var.	Load.	Var.	Load.	Var.	
1 HSPT IQ	029	001	-665	442	009	000	-162	026	031	001	.470
2 HSPT Reading Voc.	-088	008	-454	206	001	000	-102	010	-087	008	.232
3 HSPT Reading Rate	-198	039	-343	118	-043	002	511	261	-198	039	.459
4 HSPT Reading Comp.	-115	013	-524	275	-021	000	185	034	114	013	.335
5 HSPT Arithmetic	064	004	-539	291	-010	000	-090	008	062	004	.307
6 HSPT English	-036	001	-618	382	031	001	-280	078	-036	001	.463
7 HSPT Spelling	052	003	-738	545	095	009	-005	000	052	003	.560
8 HSPT Religion	-109	012	-683	466	-095	009	-310	096	-109	012	.595
9 CTP Personal Adjust.	135	018	181	033	-034	001	284	081	136	018	.151
10 CTP Social Adjust.	149	022	105	011	-142	020	062	004	150	023	.080
11 CTMM Language	086	007	-705	497	020	000	122	015	086	007	.526
12 CTMM Non-Language	070	005	-503	253	016	000	-093	009	070	005	.272
13 RTHS Religion	-105	011	-698	487	040	002	049	002	-105	011	.513
14 Grade Point Average	675	456	-542	294	096	009	-120	014	075	006	.513
15 CAT Reading Voc.	044	022	-650	423	-087	008	318	101	044	002	.536
16 CAT Reading Comp.	084	007	-664	441	023	001	076	006	084	007	.462
17 CAT Math. Reasoning	-071	005	-562	316	-195	038	-119	014	072	005	.378
18 CAT Math. Fund.	-044	002	-570	325	131	017	-168	028	045	002	.374
19 CAT Mech. of Eng.	-071	005	-718	516	104	011	-073	005	071	005	.542
20 CAT Spelling	-005	000	-726	527	006	000	022	000	004	000	.527
Total Variance	(.621)		(6.848)		(.128)		(.792)		(.172)		8.561
Per cent of Variance	7%		80%		1%		9%		2%		100%

[a]Decimals omitted.

These findings appeared to corroborate the conclusion of Wolins, MacKinney, and Stephans who found only low loadings of the grade point average on factors identified as general intelligence and science achievement as measured by standardized tests.[1] These researchers inferred that grade point average was too heavily weighted with teacher bias to warrant including it in analytical studies. Other investigators, also, have concluded that there would be minimal relationships among grade point average and standardized measures of ability and attainment.[2] Moreover, since one may assume that personality factors are involved in teacher bias, the projections from variables X_9 and X_{10} (CTP personal and social adjustment), although of minimal significance, probably contributed to the substantiation of this conclusion. Assuming, moreover, that reading rate probably contributes some portion of the variance to the broad spectrum factor identified as verbal, the projection here appeared to show its minimal importance at the ninth-grade level.

All loadings projected on Factor two (K2) were significant except for those projected by variables X_9 (CTP per-

---

[1] Supra, p. 23.

[2] Supra, p. 13.

sonal adjustment) and X_{10} (CTP social adjustment). Variables X_2 (HSPT reading vocabulary) and X_3 (HSPT reading rate) had moderate loadings of .454 and .343, respectively. Twelve variables projected high loadings: X_1 (HSPT mental ability) had a projection of .665; X_4 (HSPT reading comprehension), .524; X_5 (HSPT arithmetic), .539; X_6 (HSPT English), .618; X_8 (HSPT religion), .683; X_{12} (CTMM non-language), .503; X_{13} (RTHS religion), .698; X_{14} (grade point average), .542; X_{15} (CAT reading vocabulary), .650; X_{16} (CAT reading comprehension), .664; X_{17} (CAT mathematics reasoning), .562; and X_{18} (CAT mathematics), .570. Four loadings were very high on this Factor: X_7 (HSPT spelling) projected a loading of .738; X_{11} (CTMM language), .705; X_{19} (CAT mechanics of English and grammar), .718; and X_{20} (CAT spelling), .726. Factor two satisfied the criterion for naming a factor with a high degree of significance. The relationships represented by Factor two appeared to be a composite of verbal and mathematical ability. This Factor contributed 80% of the total variance.

As pointed out above, the ability measures contributed to Factor two in a highly significant degree. So, also, did the measures of attainment. This is substantially the same finding made in all the studies referred to in Chapter Two in which tests of ability and attainment were em-

ployed.[1] Hence, it appeared that the projections on Factor two substantiated the findings of these other studies with respect to the observed interrelationships among the factors identified as verbal and mathematical.[2] Moreover, Driscoll found that the attainment measures employed in his study contributed more to the common variance of the cognitive factor than did the ability tests.[3] In this investigation this was found to be true only of variables X_7 (HSPT spelling), X_{19} (CAT mechanics of English and grammar), and X_{20} (CAT spelling), although the two tests of attainment in religion, variables X_8 and X_{13}, did contribute more to the common variance of Factor two than did the mental ability test of the HSPT battery (variable X_1). These observations also appeared to be germane to the investigation of the structure of human abilities discussed elsewhere in this study where it was pointed out that verbal and mathematical factors have been demonstrated to be second-order factors of the factor identified as cognitive or intellective.[4]

---

[1] Supra, pp. 14-38.

[2] Supra, p. 13.

[3] Supra, p. 27.

[4] Supra, p. 12.

No significant loadings were projected on Factor three (K3). A minimal degree of significance was found in the loadings of .142 from variable X_{10} (CTP social adjustment) and of .195 from variable X_{17} (CAT mathematics reasoning). The relationships represented by this Factor might be interpreted as critical interpretation of verbal symbolization. The minimal saturations on it, however, did not satisfy the criterion for naming a factor and would suggest that identification not be attempted. This Factor accounted for 1% of the total variance. No further consideration was given to this Factor since it was assumed it could not contribute to the formulation of any meaningful conclusions.

There were five significant loadings on Factor four (K4). Variables X_6 (HSPT English) and X_9 (CTP personal adjustment) had low loadings of .280 and .284, respectively. Variables X_8 (HSPT religion) and X_{15} (CAT reading vocabulary) had moderate loadings of .310 and .318. A high loading of .511 was found on variable X_3 (HSPT reading rate). Three variables had minimally significant loadings on this Factor: X_1 (HSPT mental ability) had a loading of .162; X_4 (HSPT reading comprehension), one of .185; and X_{18} (CAT mathematics fundamentals), one of .168. Factor four satisfied the criterion for identification at a significant level. Factor

four appeared to represent abilities related to the basic skills involving verbal and numerical ability. Factor four (K4) contributed 9% of the total variance.

Most of the variables contributing to the total variance of Factor four appeared to be heavily weighted with verbal abilities. These kinds of abilities have been observed by others to contribute significantly to the basic skills factor.[1] The low but significant projection from variable X_9 (CTP personal adjustment) on this Factor appeared to corroborate further the conclusions of other researchers who have also found minimal relationships among factors identified as basic skills and personality.[2]

Factor five (K5) had no significant loadings. Minimally significant loadings were projected by three variables: X_3 (HSPT reading rate) had a loading of .198; X_9 (CTP personal adjustment), .136; and X_{10} (CTP social adjustment), .150. The abilities represented by this Factor appeared to be reactions related to such situations as speed of response. Identification of this Factor was not attempted, however, since the minimal saturations on it did not meet the criterion for naming

---

[1] Supra, p. 13.

[2] Supra, pp. 13, 15, 73.

a factor. Factor five contributed 2% of the total variance.

Variables X_3 (HSPT reading rate), X_9 (CTP personal adjustment), and X_{10} (CTP social adjustment) projected almost identical loadings on Factor five as they did on Factor one. Grade point average, however, projected a zero loading in this case whereas on Factor one this variable (X_{14}) projected a highly significant loading. Zero loadings projected by the measures of ability and zero or near zero loadings projected by the measures of attainment appeared to corroborate further the findings of other studies in which the personality factor had been observed to have only minimal relationships with intellective factors.[1]

The five extracted factors appear graphically in Figure 1 which is a profile of factor loadings on which were plotted the loadings projected by the 20 variables on each factor by the oblique rotation of the unrotated factor matrix. This figure shows the level of significance of each factor loading, also. Examination of Figure 1 leads to several observations.

Many of the loadings are not significant. Significant loadings appear only on Factors one, two, and four. Factor one has only one significant loading. Factor two has 18 sig-

---

[1] Supra, pp. 13-14, 15, 73, 79.

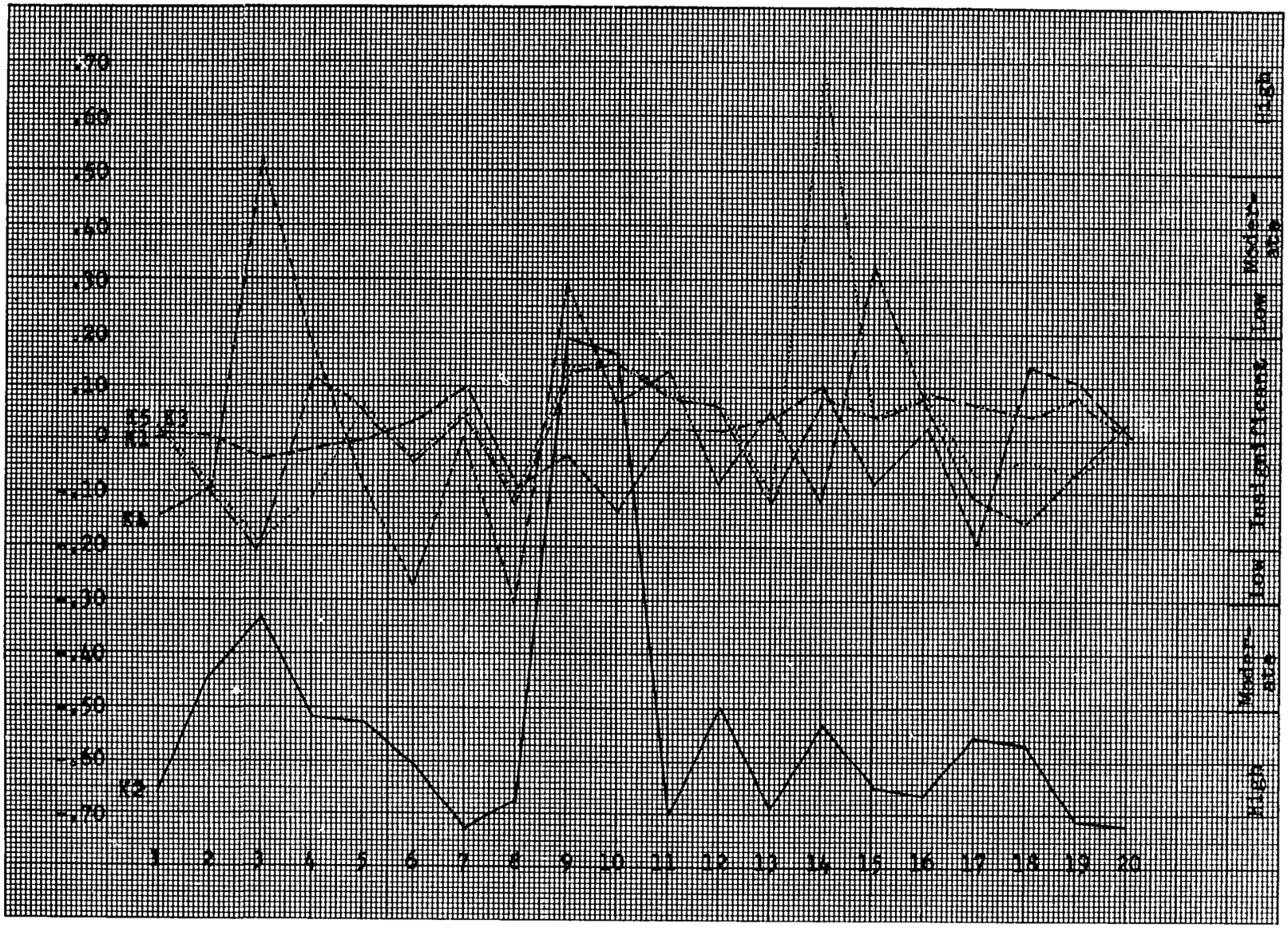


Figure 1.- Graph of rotated oblique factor loadings by variable and level of significance.

nificant loadings. Factor four has five significant loadings. More negative loadings are significant than positive. Grade point average does not correlate consistently nor highly with performance on standardized measures of ability and attainment. Personality measures project no significant loadings on any Factor. The HSPT battery projects loadings on Factor two which are significantly comparable with the projections from the CTMM and the CAT. The projections from the religion subtest of the HSPT battery and from the separate religion test (RTHS) are highly significant and almost identical on Factor two.

The purpose of this study, as stated in Chapter One, was to investigate the adequacy of the HSPT battery in general and of the religion subtest in particular. Attention is called, therefore, to the projections on the five extracted factors from the eight variables yielded by the battery. These are identified as projections from variable X_1 through X_8.

Variable X_3 (HSPT reading rate) projected a minimally significant loading of .198 on Factor one. The other seven projections on Factor one were zero or near zero. All eight variables had significant loadings on Factor two. Only zero loadings were projected on Factor three. Three variables had significant loadings on Factor four: X_3 (HSPT reading rate)

had .511; X_6 (HSPT English), .280; and X_8 (HSPT religion), .310. Moreover, variables X_1 (HSPT mental ability) and X_4 (HSPT reading comprehension) had minimally significant loadings of .162 and .185, respectively, on Factor four. Variable X_3 (HSPT reading rate) projected a loading of .198 on Factor five, precisely the same projection as it had on Factor one. Projections from the other variables on Factor five were zero or near zero.

Only Factors two and four had significant projections from the variables representing the eight subtests of the HSPT battery. Factor two was identified as a composite of verbal and mathematical ability and Factor four as basic skills involving verbal and numerical ability. The projections from the HSPT on Factor two indicated that the relationships represented by this Factor were identifiable as verbal and mathematical ability. Moreover, the projections on Factor four appeared to demonstrate that the relationships represented by this Factor were related to the basic skills involving verbal and numerical ability. It appeared, therefore, that performance on the HSPT battery did require a significant measure of verbal and mathematical ability and a significant level of attainment in verbal and numerical skills.

Comparing the factor loadings projected by the HSPT

battery with those projected by the other variables representing measures of cognitive abilities and basic skills, the HSPT appeared to contribute as substantially as the other tests to these two factors. Apparently, performance on all the subtests of the battery called for the employment of the same kinds of abilities and skills as did performance on the other tests of ability and attainment used in this study. The HSPT, therefore, can be said to measure at a significant level verbal and mathematical ability and the basic skills involving verbal and numerical ability. Moreover, it can be inferred that the battery satisfied significantly the criterion of factorial validity.

The religion subtest (variable X_8) had a highly significant loading of .683 on Factor two and a moderately significant loading of .310 on Factor four. It may be inferred, therefore, that this subtest is a measure of verbal ability, perhaps of vocabulary, within the framework of items on religion. This subtest, however, did not satisfy the criterion of factorial validity because in the factorial analysis no factor emerged to which it made a unique contribution. It appeared to be measuring the same abilities as called for in performance on the ability and attainment tests.

Summary

This chapter has presented analyses of relationships among 20 variables obtained from the performance of the sample of 59 boys on five standardized tests and the first semester grade point average in ninth-grade religion. Pearson product-moment coefficients of correlation were computed from the raw scores, arranged into a correlation matrix, and analyzed. The correlation matrix was submitted to factorial analysis using the Oblimax method, an adaptation of Thurstone's centroid method. The resulting factor loadings were arranged into an unrotated factor matrix and discussed. The unrotated factor matrix was subjected to oblique rotation on which five factors were extracted. The resulting factor loadings were arranged into a rotated factor matrix and analyzed. A profile of the five extracted factors was plotted for graphical presentation and discussed. The discussion of the obtained factors was concluded by an analysis of the HSPT battery in terms of the nature of the abilities called upon by performance on the battery in general and on the religion subtest in particular.

Conclusions

The foregoing analyses appeared to warrant a number of significant conclusions some of which corroborated the find-

ings of other studies referred to elsewhere in this study.[1] These observations, however, are made here with the suggestion that they be verified by further investigation. Hence, the following conclusions should be regarded as purely tentative:

1. Mental ability appeared to consist substantially of a composite of verbal and mathematical abilities.

2. The structure of mental ability appeared to consist of a number of overlapping, causally unrelated factors, which varied greatly in comprehensiveness.

3. The measures of ability and attainment employed in this study appeared to measure verbal and mathematical ability and basic skills related to verbal and numerical ability in a highly significant degree and about equally well.

4. The personality measures used in this study appeared to measure factors which had only minimal relationships with cognitive ability and basic verbal and numerical skills.

5. Personality factors appeared to have greater significance in the attainment of grade point average than in performance on standardized tests of attainment.

6. Grade point average appeared to be unduly weighted

---

[1] Supra, pp. 13, 37-38.

with teacher bias and should, therefore, probably not be included in studies of factorial analysis where attainment and basic skills are concerned.

7. The HSPT battery appeared to have highly significant concurrent validity. To the extent that the other tests of ability and attainment used in this study have factorial validity, the HSPT may be inferred to meet this criterion.

8. The religion subtest (variable X_8) of the HSPT appeared to measure verbal ability within the framework of items on religion. It appeared to be measuring the same abilities as required for performance on the ability and attainment tests. It did not satisfy the criterion of factorial validity.

9. Further investigation of the HSPT battery in general and of the religion subtest in particular is recommended.

CHAPTER V

ITEM ANALYSIS OF THE RELIGION SUBTEST OF THE HSPT BATTERY

Introduction

The religion subtest (variable X_8) of the HSPT battery is test number 14 of the battery. This test consists of 55 multiple-choice items. The items sample the major objectives included in the Baltimore Catechism. The Creed, the Commandments, and the Sacraments are about equally represented in the test items. Each item has four responses from which the correct response is to be selected. The number right constitutes the raw score. The split-half reliability estimate is reported in Table 2 as 0.75 with a sigma of 8.40 and a standard error of measurement of 4.20. For the sample of 59 boys, the mean was 33.88 with a sigma of 10.38 and a standard error of the mean of 1.36 as reported in Table 5.

Ahmann and Glock point out that the procedure known as "item analysis" consists in "re-examining each item of a

test for the purpose of discovering its strengths and flaws."[1] The technique usually "concentrates on two vital features of each test item: its level of difficulty and its discriminating power."[2] The level of difficulty value indicates the percentage of examinees who answer each test item correctly. The discriminating power value indicates the ability of each test item to differentiate between examinees who have performed well and those who have performed poorly.

Level of Difficulty and Discriminating Power

Since the value of the discriminating power of each test item is determined with maximum significance by using the proportions of successes in the highest and lowest 27% of the sample, Davis recommends determining the level of difficulty by using the same two groups. Moreover, he recommends a formula which includes a correction for chance and which takes into consideration the number of examinees who do not reach a given test item.[3] This procedure yields two

---

[1]J. Stanley Ahmann and Marvin D. Glock, Evaluating Pupil Growth (Boston: Allyn and Bacon, 1958), p. 292.

[2]Ibid.

[3]Davis, op. cit., pp. 30-34.

proportions for each test item which are entered in Davis' item-analysis chart to determine the level of difficulty and the discriminating power for each item. The formula, therefore, is:

For the highest 27% of the sample:

$$P_h = \frac{R_h - \frac{W_h}{K - 1}}{N_h - NR_h}$$

For the lowest 27% of the sample:

$$P_l = \frac{R_l - \frac{W_l}{K - 1}}{N_l - NR_l}$$

where: N_h = the number of testees in the highest 27% of the sample,
R_h = the number of testees in the highest 27% of the sample that answer the item correctly,
W_h = the number of testees in the highest 27% of the sample that answer the item incorrectly,
NR_h = the number of testees in the highest 27% of the sample that do not reach the item in the time limit,
N_l = the number of testees in the lowest 27% of the sample,
R_l = the number of testees in the lowest 27% of the sample that answer the item correctly,
W_l = the number of testees in the lowest 27% of the sample that answer the item incorrectly,
NR_l = the number of testees in the lowest 27% of the sample that do not reach the item in the time limit,
K = the number of choices in the item.

In order to apply this formula, it was necessary to determine the highest and lowest 27% of the sample. Since 27% of the total sample of 59 boys was found to be 15.93, this product was rounded to 16. The criterion of internal-consistency was used to determine the constituents of the two groups, i.e. the obtained scores representing the performance of the sample on the total religion subtest (variable X_8) were arranged in rank order and the highest and lowest 16 were removed for the determination of the level of difficulty (P) and the discriminating power (D) of the individual items while the middle 46% was discarded.

Using the obtained scores of the highest and lowest 27%, a cumulative frequency distribution of subjects not reading each item was constructed for each group as reported in Table 9. Moreover, a frequency distribution was constructed to determine the number in the highest and lowest 27% that marked the correct answer to each item and the number that marked any wrong answer. The percentage that marked the correct answer was computed by group for each item. These data are included in Table 9. Utilizing Davis' item-analysis chart, the level of difficulty and the discriminating power of each item were found.[1] The values for level of difficulty

---

[1] Ibid., p. 43.

TABLE 9

NUMBER IN HIGHEST AND LOWEST 27% OF THE STUDENTS WHO MARKED THE CORRECT ANSWER, THE INCORRECT ANSWER, AND WHO DID NOT REACH THE ITEM IN THE TIME LIMIT WITH THE DIFFICULTY AND DISCRIMINATION INDICES CALCULATED

Item	Highest 27%				Lowest 27%				P	D	Item	Highest 27%				Lowest 27%				P	D
	R	W	NR	P	R	W	NR	P				R	W	NR	P	R	W	NR	P		
1	16	0	0	97-	15	1	0	92	.84	.11	29	14	2	0	83	8	6	2	43	.57	.27
2	16	0	0	97-	15	1	0	92	.84	.11	30	13	3	0	75	0	14	2	4+	.47	.43
3	14	2	0	83	10	6	0	50	.59	.24	31	14	2	0	83	4	9	3	8	.47	.57
4	13	3	0	75	7	9	0	25	.50	.33	32	7	9	0	25	3	10	3	4+	.27	.26
5	15	1	0	92	12	4	0	67	.67	.24	33	15	1	0	92	4	9	3	8	.50	.67
6	13	3	0	75	6	10	0	17	.48	.40	34	10	6	0	50	6	6	4	33	.45	.11
7	8	8	0	33	11	5	0	58	.47	-.16	35	12	4	0	67	6	6	4	33	.50	.22
8	16	0	0	97-	14	2	0	83	.76	.23	36	11	5	0	58	3	9	4	0	.39	.60
9	13	3	0	75	11	5	0	58	.59	.11	37	16	0	0	97-	9	2	5	76	.73	.29
10	14	2	0	83	6	10	0	17	.50	.47	38	11	5	0	58	3	8	5	3	.39	.50
11	15	1	0	92	8	8	0	33	.57	.46	39	11	5	0	58	3	8	5	3	.39	.50
12	14	2	0	83	8	8	0	33	.54	.33	40	11	5	0	58	5	6	5	25	.46	.20
13	16	0	0	97-	7	8	1	29	.56	.59	41	12	4	0	67	3	8	5	3	.42	.56
14	15	1	0	92	12	3	1	73	.69	.20	42	14	2	0	83	4	6	6	20	.51	.44
15	13	3	0	75	4	11	1	2	.44	.64	43	13	3	0	75	2	6	8	0	.43	.70
16	14	2	0	83	6	9	1	20	.51	.44	44	11	5	0	58	3	5	8	17	.43	.29
17	16	0	0	97-	12	3	1	73	.71	.31	45	11	5	0	58	3	5	8	17	.43	.29
18	13	3	0	75	11	4	1	64	.60	.07	46	12	4	0	67	2	6	8	0	.41	.65
19	14	2	0	83	6	9	1	20	.51	.44	47	15	1	0	92	6	2	8	67	.67	.24
20	13	3	0	75	4	11	1	2	.44	.64	48	14	2	0	83	2	6	8	0	.45	.77
21	15	1	0	92	4	11	1	2	.48	.80	49	15	1	0	92	3	5	8	17	.52	.57
22	13	3	0	75	3	12	1	3+	.44	.60	50	5	11	0	8	0	8	8	6+	.19	.04
23	15	1	0	92	5	9	2	14	.52	.60	51	10	6	0	50	0	8	8	6+	.38	.38
24	15	1	0	92	6	8	2	24	.54	.51	52	15	1	0	92	2	6	8	0	.48	.86
25	6	10	0	17	2	12	2	4+	.23	.20	53	16	0	0	97-	5	2	9	62	.67	.39
26	15	1	0	92	9	5	2	52	.62	.34	54	16	0	0	97-	5	2	9	62	.67	.39
27	15	1	0	92	8	6	2	43	.59	.39	55	15	1	0	92	4	3	9	43	.59	.39
28	13	3	0	75	6	8	2	24	.49	.34											

and discriminating power of each item are reported in Table 9, also.

Upon entering the proportions in the item-analysis chart, it was observed that six items in the lowest 27% group had minus proportions. These minus values resulted from a disproportionately high number of incorrect answers. There were no minus values in the highest 27% group. According to Davis, "there is no straightforward way of using the item-analysis chart"[1] when the signs of the two proportions for any given item are different. Nevertheless, Davis presents an adaptation of the procedure by which to determine the percentage of successes in this case.

> Practical experience has convinced the writer that reasonably serviceable discrimination indices can be obtained for items when the percentage of successes in the lowest 27% of the sample is smaller than 1 by altering the percentage arbitrarily to the percentage that would be obtained if one-half a testee had answered the item correctly. This percentage is determined by the number of testees that tried the item . . . The percentage is 1 if the numerical value of the denominator . . . is 34 or more . . . However, percentages corresponding to as small a number of testees as 5 are included. . . . A discrimination index obtained by this procedure is always an underestimate, and this fact can be noted conveniently by adding a plus-sign superscript to the index.[2]

---

[1] Ibid., p. 34.

[2] Ibid., p. 35.

Applying Davis' suggestion, the percentage of successes on items 22, 25, 30, 32, 50, and 51 for the lowest 27% was altered, and each is indicated by a plus-sign superscript in Table 9.

Several cases in which it was found necessary to alter the percentage of successes on an item in the highest 27% group occurred in items 1, 2, 8, 13, 17, 37, 53, and 54 where the percentage of successes was 100%. In this case Davis recommends that the percentage be altered "to the percentage that would be obtained if one-half a testee had answered the item incorrectly."[1] Davis presents a table by which this value may be obtained depending on the number of testees that read the item. Since N = 16 in all the cases under consideration, the P for the forementioned items was found to be 97. Entries in Table 9 that have been altered in this manner are indicated by a minus-sign superscript.

The criterion for significance of the level of difficulty value is discussed by Guilford.

> Remembering that variance means individual differences and the greater the variance, the more we have dispersed individuals in measurement, it can be stated that the item that will produce the greatest dispersion is of median difficulty. It is an item passed by half of the group and failed by half of the group. . . .

---

[1] Ibid., p. 36.

> Items of moderate difficulty . . . yield the maximum variance.[1]

Ahmann and Glock, also, discuss the criterion for significance of the level of difficulty value.[2] They submit that items in achievement tests which fall within a range of 40 per cent to 70 per cent of difficulty display more discriminating power than items with widely varying levels of difficulty.

Examination of Table 9 suggested several observations relevant to the level of difficulty of the items in the religion subtest of the HSPT battery. The range of the proportions was from 19% to 84% inclusive. Five items had a P value which exceeded 70% and seven items had a P value which was less than 40%. In both cases over half of these items were very close to the limits of the range of significance. Apparently, some five items (1, 2, 25, 32, and 50) may be considered as not having a satisfactory level of difficulty to warrant their use in this test.

Ahmann and Glock discuss the criterion for significance of the discriminating power value.

> The maximum size of the index is +1.00 and the minimum size is -1.00. In the first case maximum discrimination

---

[1]Guilford, op. cit., p. 450.

[2]Ahmann and Glock, op. cit., pp. 297-98.

> occurs in the desired direction; in the second case the discrimination occurs in the direction exactly opposite to that desired. Any negative value means that the test item discriminates to some degree in the wrong direction. Hence, the discriminating power of the test item is unsatisfactory. Positive values indicate that the test item discriminates in the desired direction, even though the test item may not be completely satisfactory in this respect. The larger the positive value the better. It is difficult to establish a suitable minimum positive value below which the discriminating power of a test item is considered faulty. Certainly values less than +0.20 indicate that the discriminating power of the test item is questionable.[1]

Returning to Table 9, it was evident that seven test items fell below +0.20 in discriminating power. These items were 1, 2, 7, 9, 18, 34, and 50. Attention is called to the fact that items 1, 2, and 50 also failed to satisfy the criterion for level of difficulty.

Item 7 had a D of -0.16 and, although its P = .47, should definitely be eliminated from the test or revised. More of the examinees in the lowest 27% answered this item correctly than in the highest 27%. It certainly did not discriminate in the desired direction. Items 1 and 2 each had a D of +0.11 and a P of .84. Sixteen examinees in the highest 27% and 15 in the lowest 27% answered these items correctly. Neither item discriminated satisfactorily. Both items should be revised or eliminated from the test. Item 9

---

[1] Ibid., p. 297.

for which D = +0.11 and P = .59 appeared to be a very weak item. Although all examinees in both groups tried it, two more failed it in the lowest group than in the highest group. This item did not discriminate satisfactorily. It should be eliminated or revised. Item 18 had a D of +0.07 and a P of .60. This item was very low in discriminating power. Two more examinees in the highest group answered it correctly than in the lowest group one of whom did not reach it. This item did not discriminate satisfactorily. It should be revised or eliminated from the test. Item 34 had a D of +0.11 and a P of .45. Ten examinees in the highest group answered this item correctly and six failed it. In the lowest group six answered it correctly, six failed it, and four did not reach it. This item did not discriminate satisfactorily. It should be eliminated or revised. Item 50 for which D = +0.04 and P = .19 appeared to be too difficult. Five examinees in the highest 27% group answered it correctly and 11 failed it. In the lowest 27% group no one answered it correctly, eight failed it, and eight did not reach it. This item satisfied neither criterion of significance and should, therefore, be revised or eliminated from the test.

Items 25 and 32, as noted above, fell below the lower limit of the range of the criterion for significance of the

level of difficulty. Item 25 had a D of +0.20 and a P of .23. In the highest group six examinees answered it correctly and 10 failed it. Only two in the lowest group answered it correctly while 12 failed it and four did not reach it. This item appeared to be too difficult. Moreover, it did not discriminate satisfactorily. It should be revised or eliminated. Item 32 had a D of +0.26 and a P of .27. In the highest group seven examinees answered it correctly and nine failed it. In the lowest group three answered it correctly, 10 failed it, and three did not reach it. This item appeared to be too difficult and should be revised or eliminated from the test.

No problem appeared to arise with respect to the remaining 46 items. All of these items satisfied the criteria for significance of level of difficulty and discriminating power.

The two values of the 55 test items are plotted graphically in Figure 2. The rectangular area in the center of this figure sets the limits for significance of level of difficulty and discriminating power. All items falling beyond the limits of this domain, as pointed out above, should be revised or eliminated from the test. Examination of this figure indicated that the nine items discussed above did not

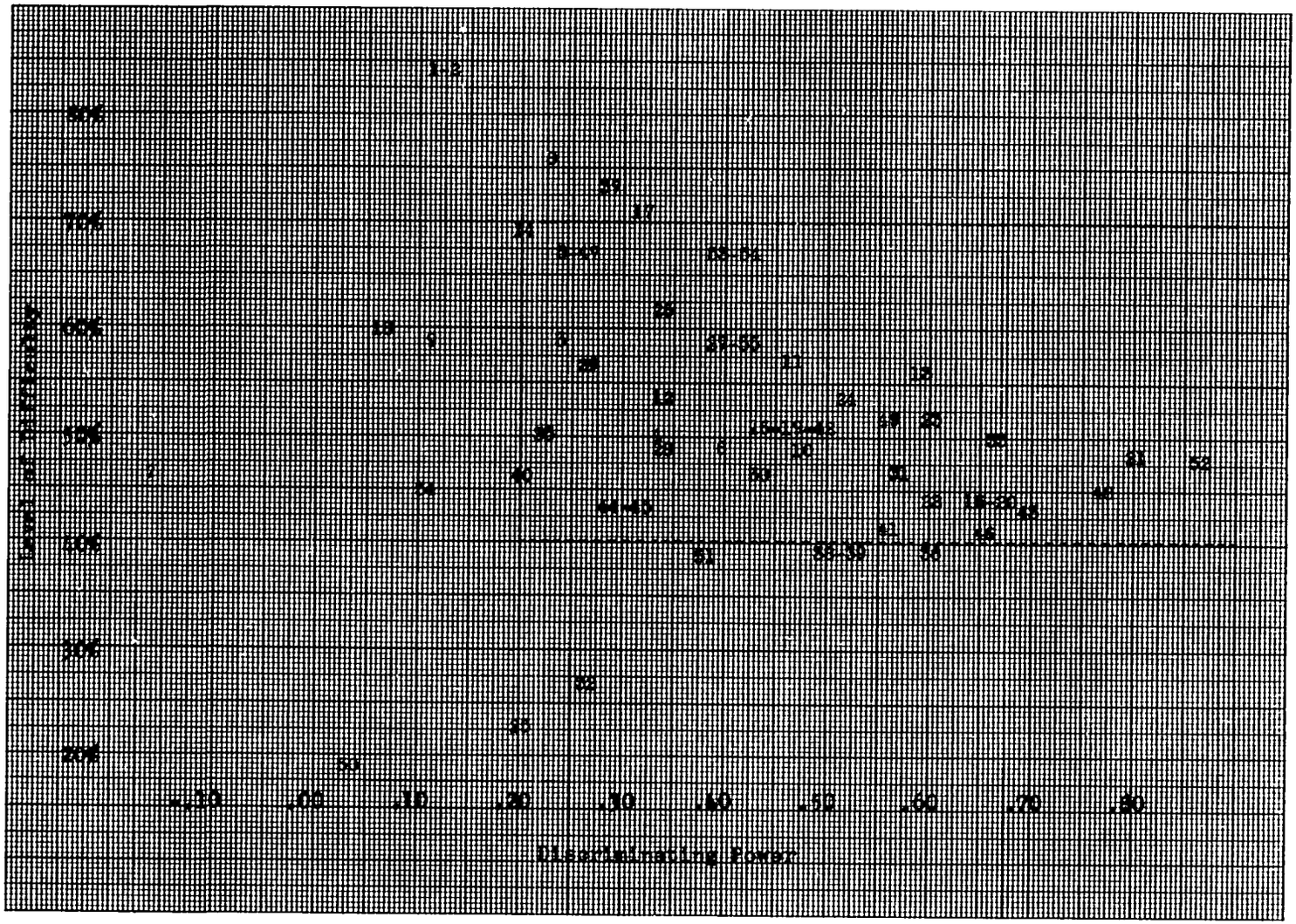


Figure 2.- Scattergram indicating significance by levels of difficulty and discrimination of 55 items, by item number, in HSPT religion subtest. Items in framed area are of adequate levels of difficulty and discrimination to be retained in the subtest.

fall within the boundaries of significance.

These items are:

1. Priests can forgive sins. So can

   a. all religious. b. all Catholics. c. nuns. d. bishops.

2. A person cannot be a member of the Church unless he is

   a. confirmed. b. baptized. c. ordained. d. married.

7. Marie borrows a pen from Louis for one week. Three weeks later she still has not returned it. When Louis asks for the pen, Marie

   a. may lawfully claim the pen as her own.
   b. must return the pen at once.
   c. must return the pen and pay Louis two weeks' rental.
   d. must return the pen and confess her sin the next time she goes to confession.

9. Under no circumstances can we be forgiven a sin without

   a. confessing it. b. examining our conscience. c. doing penance. d. being sorry for it.

18. The words "There is nothing God cannot do" mean that God is

   a. almighty. b. all-present. c. supreme. d. infinite.

25. God can never change because He is

   a. eternal. b. supreme. c. infinite. d. spiritual.

32. The eternal punishment due us on account of our mortal sins is taken away at the time of

a. contrition. b. confession. c. absolution. d. penance.

34. Ray's father deliberately kills a pig on his farm. He is

a. guilty of murder. b. not sinning at all. c. hurting defenseless animals. d. breaking God's law.

50. Linda wants Shirley's pen so badly that she plans how she may steal it. Linda is offending against the

a. tenth commandment. b. seventh commandment. c. fourth commandment. d. first commandment.[1]

Summary, Conclusions, and Recommendations

This chapter has been dedicated to an item-analysis of the religion subtest (variable X_8) of the HSPT. A statement was made as to the nature of this procedure. Values for level of difficulty and discriminating power were computed using Davis' formula and item-analysis chart. This procedure was presented by means of a table, a figure, criteria of significance, and discussion. An evaluation of the 55 items in the test was made in terms of their respective levels of difficulty and discriminating power. Nine items were considered to be faulty, and it was recommended that they be eliminated from the test or revised. The remaining 46 items were demon-

---

[1]Anderhalter, Colestock, and Gawkowski, High School Placement Test, pp. 19-23.

strated to satisfy the criteria for significance of level of difficulty and discriminating power.

This investigation demonstrated that the religion subtest (variable X_8) of the HSPT appeared to be faulty in that it contained nine items which were in need of revision. These nine items represented slightly more than 16 per cent of the total number of items in the subtest. While this did not appear to be a high percentage of the total number of items, it was significant enough to warrant the recommendations for revision made above. Moreover, since it is recognized that both the homogeneity and the smallness of the size of the sample used in the present study may have influenced the findings in this analysis, it is recommended that, after revision, the religion subtest be investigated further using a randomized and larger sample.

CHAPTER VI

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Summary

The administration of beginning high school placement tests has long been an integral part of high school guidance programs. Batteries of tests designed for this purpose are more recent. Until 1955 there was no such battery which comprised tests that were designed by the same experts and standardized on Catholic school students and which included a test on religious education. Such a test, High School Placement Test, Battery for Entering Freshmen, (Form A), was published in 1955 and soon found widespread popularity among Catholic school guidance personnel.

The HSPT, however, appeared to be unsatisfactory after some years of utilization because it apparently did not meet the criteria of reliability and validity customarily expected of such instruments. Although the use of the test continued, there appeared to be a widespread attitude of lack of confidence in it. Some experts and many test users suggested that

its use be discontinued.

This study was designed to attack the problem posed by these circumstances by determining the adequacy of the battery in general and of the religion subtest in particular. It was decided to attempt to achieve this objective by investigating the factorial validity of the battery and by analyzing the religion subtest by factorial analysis and by item analysis. In order to accomplish this, it was necessary to isolate and compare the factors extracted from the original data which were obtained from the administration of five standardized tests of ability, attainment, and personality and from the first semester grade point average in ninth-grade religion. The tests used in this study were: High School Placement Test (HSPT), California Achievement Tests (CAT), California Short-Form Test of Mental Maturity (CTMM), California Test of Personality (CTP), and Religion Test for High Schools (RTHS).

The sample consisted of 59 male beginning ninth-grade students. These subjects were a substantially homogeneous group in terms of age, elementary education in Catholic parochial schools, socio-economic background, and religious background.

Although the tests yielded a total of 40 subtest raw

scores, only subtotal raw scores were used so that 19 variables represented the five standardized tests used in the study. An additional variable was the first semester grade point average in ninth-grade religion.

Pearson product-moment coefficients of correlation were determined for these 20 variables and arranged into a correlation matrix. The correlation matrix was submitted to factorial analysis by the Oblimax method, a modification of Thurstone's centroid method adapted to IBM processing. The five extracted factors obtained by this procedure were arranged into an unrotated factor matrix. The variance of each loading on each factor, the communalities of the variances, the total variance of each factor, and the per cent of variance of each factor were computed.

Criteria for the number of factors to be extracted and for the determination of the significance of factors were discussed and utilized in the examination of the unrotated factor matrix. Preliminary observations were made concerning the implications of this matrix.

The unrotated factor matrix was submitted to oblique rotation by the Oblimax method. Five factors were extracted and the new factor loadings were arranged into a rotated factor matrix.

Factor one (K1) revealed relationships among four loadings which appeared to show it had nothing in common with the basic skills tests and very little with the personality measures used in this study. This Factor, however, was not identified because it did not satisfy Thurstone's criterion for naming a factor.

Factor two (K2) revealed highly significant relationships among all the loadings except those projected by the personality measures (variables X_9 and X_{10}). These relationships appeared to be a composite of verbal and mathematical ability.

Factor three (K3) had no significant loadings. Two minimally significant loadings indicated the relationships represented by Factor three might be interpreted as critical interpretation of verbal symbolization. Since this Factor did not satisfy the criterion for naming a factor, no attempt was made to identify it.

Factor four (K4) revealed five significant and three minimally significant loadings. This Factor satisfied the criterion for naming a factor. The relationships revealed by Factor four appeared to represent abilities related to the basic skills involving verbal and numerical ability.

Factor five (K5) had no significant loadings. Three

minimally significant projections suggested the relationships revealed by this Factor appeared to be reactions to such situations as speed of response. Identification of this Factor was not attempted since it did not satisfy the criterion for naming a factor.

Attention was called to the implications of the analysis of the five obtained factors and their interrelationships with respect to the HSPT in general and the religion subtest in particular. It was observed that performance on the HSPT called for substantially the same kinds of abilities and skills as did performance on the other tests of ability and attainment used in this study. It was further observed that these same findings were germane to the religion subtest.

The religion subtest (variable X_8) was subjected to an item analysis according to Davis' method. The level of difficulty and the discriminating power of each of the 55 items of which the test consists were determined. Nine test items were found to be faulty, and it was recommended, therefore, that they should be removed from the test or revised. The remaining 46 items met the criteria for significance, and it was recommended, therefore, that they should be retained.

Conclusions and Recommendations

The aim of this study was to determine the adequacy of

the HSPT in general and of the religion subtest (variable X_8) in particular. The analyses made in this study warranted a number of conclusions and recommendations.

1. Mental ability appeared to consist substantially of a composite of verbal and mathematical abilities.

2. The structure of mental ability appeared to consist of a number of overlapping, causally unrelated factors, which varied greatly in comprehensiveness.

3. The measures of ability and attainment employed in this study appeared to measure verbal and mathematical ability and basic skills related to verbal and numerical ability in a highly significant degree and about equally well.

4. The personality measures used in this study appeared to measure factors which had only minimal relationships with cognitive ability and basic verbal and numerical skills.

5. Personality factors appeared to have greater significance in the attainment of grade point average than in performance on standardized tests of attainment.

6. Grade point average appeared to be unduly weighted with teacher bias and should, therefore, probably not be included in studies of factorial analysis where attainment and basic skills are concerned.

7. The HSPT battery appeared to have highly significant

concurrent validity. To the extent that the other tests of ability and attainment used in this study have factorial validity, the HSPT may be inferred to meet this criterion and should, therefore, be considered as an adequate instrument for measuring ability and basic skills.

8. The religion subtest (variable X_8) of the HSPT appeared to measure verbal ability within the framework of items on religion. It appeared to be measuring the same abilities as required for performance on the ability and attainment tests. It did not satisfy the criterion of factorial validity.

9. The items of the religion subtest (variable X_8) of the HSPT met the criteria for level of difficulty and discriminating power except in the case of nine items which should be eliminated from the subtest or revised.

10. Since the findings in this investigation may have resulted from the homogeneity and the smallness of the size of the sample in the present study, it is recommended that the HSPT in general and the religion subtest (variable X_8) in particular be submitted to further investigation. A replication of this study with revised items and a larger sample of students in other Catholic schools representative of various geographical areas and socio-economic levels is, therefore, recommended.

REFERENCES

Books

Adcock, C. J. Factorial Analysis for Non-Mathematicians. London: Cambridge University Press, 1954.

Ahmann, J. Stanley, and Glock, Marvin D. Evaluating Pupil Growth. Boston: Allyn and Bacon, 1958.

Anastasi, Anne. Psychological Testing. New York: The Macmillan Co., 1954.

Buros, Oscar Krisen (ed.). The Fifth Mental Measurements Yearbook. Highland Park, New Jersey: The Gryphon Press, 1959.

Cattell, R. B. Description and Measurement of Personality. Yonkers-on-Hudson, New York: World Book Co., 1946.

Cronbach, Lee J. Essentials of Psychological Testing. New York: Harper and Brothers, 1949.

Davis, Frederick B. Item-Analysis Data. Cambridge, Mass.: Graduate School of Education, Harvard University, 1949.

Doyle, Andrew M. Some Aspects of Ability and Achievement in High School Girls. Washington, D.C.: The Catholic University of America Press, 1952.

Driscoll, Justin A. Factors in Intelligence and Achievement. Washington, D.C.: The Catholic University of America Press, 1952.

French, John W. The Description of Aptitude and Achievement Tests in Terms of Rotated Factors. ("Psychometric Monographs," No. 5.) Chicago: University of Chicago Press, 1951.

French, John W. The Description of Personality Measurements in Terms of Rotated Factors. Princeton: Educational Testing Service, 1953.

Fruchter, Benjamin. Introduction to Factor Analysis. New York: D. Van Nostrand Co., 1954.

Guilford, J. P. Fundamental Statistics in Psychology and Education. New York: McGraw-Hill Book Co., 1956.

_____. Psychometric Methods. New York: McGraw-Hill Book Co., 1936.

Haller, A. O., and Miller, Irwin W. The Occupational Aspiration Scale: Theory, Structure and Correlates. ("Department of Sociology and Anthropology Technical Bulletin 288.") East Lansing, Michigan: Michigan State University Cooperative Extension Service, 1963.

Lindquist, E. F. (ed.). Educational Measurement. Washington, D.C.: American Council on Education, 1959.

McCormick, Sister William Pauline. Factors of Intelligence in High and Low Cognitive Ability Groups. Washington, D.C.: The Catholic University of America Press, 1954.

Roesslein, Charles G. Differential Patterns of Intelligence Traits Between High Achieving and Low Achieving High School Boys. Washington, D.C.: The Catholic University of America Press, 1953.

Russell, David H. The Dimensions of Children's Meaning Vocabularies in Grades Four through Twelve. ("University of California Publications in Education," No. 11.) Berkeley: University of California Press, 1954.

Ruszel, Humphrey. Test Patterns in Intelligence. Washington, D.C.: The Catholic University of America Press, 1952.

Spearman, Charles. The Abilities of Man. London: Macmillan, 1927.

Spearman, Charles, and Jones, LL. Wynn. Human Ability. London: Macmillan & Co., Ltd., 1951.

Thomson, Godfrey H. The Factorial Analysis of Human Ability. 5th ed. Boston: Houghton Mifflin Co., 1960.

Thurstone, L. L. Multiple Factor Analysis. Chicago: University of Chicago Press, 1947.

_____. Primary Mental Abilities. ("Psychometric Monographs," No. 1.) Chicago: University of Chicago Press, 1938.

Vernon, P. E. The Structure of Human Abilities. London: Methuen and Co., Ltd., 1950.

Wolfle, Dael. Factor Analysis to 1940. ("Psychometric Monographs," No. 3.) Chicago: University of Chicago Press, 1940.

Yela, Mariano. Psicologia de las aptitudes. "Biblioteca Hispanica de Filosofia." Madrid: Editorial Gredos, 1956.

Articles and Periodicals

Carroll, J. B., and Schweiker, R. F. "Factor Analysis in Educational Research," Review of Educational Research, XXI (1951), 368-88.

Cottle, William C. "High School Placement Test," The Fifth Mental Measurements Yearbook, ed. Oscar Krisen Buros. Highland Park, New Jersey: The Gryphon Press, 1959.

Dodd, S. C. "The Theory of Factors," Psychological Review, XXXV (1928), 211-31; 261-79.

Jones, Robert A. "High School Placement Test," The Fifth Mental Measurements Yearbook, ed. Oscar Krisen Buros. Highland Park, New Jersey: The Gryphon Press, 1959.

Rupiper, Omer J. "Multiple Factor Analysis of Academic Achievement: A Comparative Study of Full-Blooded Indian and White Children," Journal of Experimental Education, 28 (1960), 177-205.

Wolins, Leroy, MacKinney, A. C., and Stephans, Paul. "Factor Analyses of High School Science Achievement Measures," Journal of Educational Research, 54 (1961), 173-77.

Reports

California Test Bureau. Basic Data for Factor Analytic Studies with the California Test of Mental Maturity Series, 1963 Revision. Monterey, California: California Test Bureau, 1963.

California Test Bureau. California Test of Personality. "Summary of Investigations, Number One." Monterey, California: California Test Bureau, 1950.

California Test Bureau. 1957 Technical Report on the California Achievement Tests. Monterey, California: California Test Bureau, 1957.

California Test Bureau. 1957 Technical Report on the California Test of Mental Maturity. Los Angeles: California Test Bureau, 1957.

Scholastic Testing Service, Inc. STS High School Placement Test Technical Report. Bensenville, Illinois: Scholastic Testing Service, Inc., 1963.

Unpublished Material

Anderson, Harry E., Jr., and Leton, Donald A. "Factor Analyses of the California Test of Mental Maturity." An unpublished paper prepared for publication in a professional journal. Los Angeles, 1963. (Mimeographed.)

Caffrey, John, and Smith, Thomas W. "Preliminary Identification of Some Factors in the Davis-Eells Games." An unpublished paper read before the Annual Convention of the American Psychological Association, San Francisco, September 6, 1955.

Elias, Jack Z. "Non-Intellective Factors in Certain Intelligence and Achievement Tests: An Analysis of Factors in Addition to the Cognitive Entering into the Intelligence and Achievement Scores of Children at the Sixth Grade Level." Unpublished Doctor's dissertation, New York University, 1951.

Spivey, Gordon M. "The Relationship between Temperament and Achievement of a Selected Group of John Muir College

Students." Unpublished Doctor's dissertation, University of Southern California, Los Angeles, 1950.

Other Sources

Anderhalter, Oliver F., Colestock, Ruth, and Gawkowski, R. Stephen. High School Placement Test, Battery for Entering Freshmen, (Form A). Chicago: Scholastic Testing Service, Inc., 1955.

_____. "High School Placement Test, Battery for Entering Freshmen (Form A)." Scholastic Tests Manual, Scholastic Testing Service, Inc., Chicago, 1955.

Clark, W. W., and Tiegs, E. W. California Short-Form Test of Mental Maturity, Grades 9 to 13, Secondary, 1957 S-Form, Form A. Los Angeles: California Test Bureau, 1957.

_____. "California Short-Form Test of Mental Maturity, Grades 9 to 13, Secondary, 1957 S-Form, Form A." California Test Bureau Manual, California Test Bureau, Los Angeles, 1957.

Schmidt, A. G., and Anderhalter, O. F. Religion Test for High Schools, Form B. Chicago: Loyola University Press, 1955.

_____. "Religion Test for High Schools." General Manual, Loyola University Press, Chicago, 1955.

Thorpe, L. P., Clark, W. W., and Tiegs, E. W. California Test of Personality, 1953 Revision, Secondary, Grades 9 to College, Form AA. Los Angeles: California Test Bureau, 1953.

_____. "California Test of Personality, 1953 Revision, Secondary, Grades 9 to College, Form AA." California Test Bureau Manual, Los Angeles, 1953.

Tiegs, E. W., and Clark, W. W. California Achievement Tests Complete Battery, Advanced, Grades 9 to 14, Form AA. Los Angeles: California Test Bureau, 1950.

Tiegs, E. W., and Clark, W. W. "California Achievement Tests Complete Battery, Advanced, Grades 9 to 14, Form AA." California Test Bureau Manual, California Test Bureau, Los Angeles, 1951.