Spring 1-1977

A STUDY OF FACTOR B OF THE EARLY SCHOOL PERSONALITY QUESTIONNAIRE(ESPQ) AS A VALID MEASURE OF GENERAL MENTAL ABILITY

Claude S. Iles
Kansas State College Of Pittsburg

A STUDY OF FACTOR B OF THE EARLY SCHOOL PERSONALITY QUESTIONNAIRE(ESPQ) AS A VALID MEASURE OF GENERAL MENTAL ABILITY

A Thesis Submitted to the Graduate Division in Partial Fulfillment of the Requirements for the Degree of Master of Science

By

Claude S. Iles

KANSAS STATE COLLEGE OF PITTSBURG

Pittsburg, Kansas

January, 1977

ABSTRACT

This study investigated the validity of the general ability factor, Factor B of the Early School Personality Questionnaire(ESPQ), in two ways. First, validity coefficients for Factor B were obtained by computing product-moment correlation coefficients between the ESPQ Factor B raw scores of 40 developmentally disabled subjects and their raw scores on the Wechsler Intelligence Scale for Children-Revised (WISC-R). Second, the Kruskal-Wallis rank sums test and a distribution-free multiple comparisons test were employed to determine if there were significant differences among the mean ESPQ Factor B raw scores of subjects designated as mentally retarded, learning disabled, or normal. Specifically, it was tested if there were significant differences between the mean ESPQ Factor B raw scores of (1) the mentally retarded subjects and the learning disabled subjects, (2) the mentally retarded subjects and the normal subjects, and (3) the learning disabled and the normal subjects.

In the first case, all of the validity coefficients obtained for Factor B of the ESPQ were significant at the .001 level with the exception of the validity coefficient obtained from the correlation of the ESPQ Factor B raw scores with the raw scores of the WISC-R Coding subtest. It was concluded that the ESPQ Factor B raw scores, to a degree, measured the general mental ability of the developmentally disabled subjects much like the raw scores of the WISC-R

measured the general mental ability of the developmentally disabled subjects employed in the study.

In the second case, the results indicated that there were significant differences at the .001 level among the mean ESPQ Factor B raw scores of the mentally retarded, learning disabled, and normal subjects. The mean ESPQ Factor B raw scores of the learning disabled subjects and the normal subjects were found to be significantly higher than the mean ESPQ Factor B raw scores of the mentally retarded subjects, and (2) the mean ESPQ Factor B raw scores of the learning disabled subjects and the normal subjects did not significantly differ. Respectively, it was concluded that (1) the learning disabled and normal subjects were higher than the mentally retarded subjects in mean level of general mental ability as measured by the ESPQ Factor B raw scores, and (2) the learning disabled subjects and the normal subjects were about equivalent in mean level of general mental ability as measured by the ESPQ Factor B raw scores. The latter finding was interpreted as agreeing with a generally accepted criterion for diagnosing learning disabilities in that learning disabled children are generally differentiated from normal children on the basis of specific differences in basis psychological processes rather than on the basis of differences in general mental ability. It was concluded that the ESPQ Factor B raw scores apparently were not an accurate measure of those specific aspects of psychological processes which differentiated the learning disabled child from the

normal child in the study.

The general findings of the study suggested that the ESPQ Factor B raw scores, to a good degree, were a valid measure of the general mental ability of the subjects employed in the study since (1) the validity coefficients obtained for Factor B were generally substantial with the sample of developmentally disabled subjects employed, and (2) significant differences in mean ESPQ Factor B raw scores were measured among the subjects designated as mentally retarded, learning disabled, or normal.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

The experimental use of the Early School Personality Questionnaire(ESPQ) at the Joplin Regional Center for Developmentally Disabled, Joplin Missouri, led the investigator to question if Factor B on the instrument, an intelligence factor, was a valid measure of general mental ability for the population of children seen at the center. The authors of the ESPQ (Coan & Cattell, 1972, p. 6) stated that "the child who scores high on Factor B tends to be 'bright' and abstract-thinking, while the low-scoring child is more concrete-thinking. The intelligence factor is simply a rapid screening measure which allows the classroom teachers to assess general ability."

To the knowledge of the investigator, there have been no studies which attempt to evaluate the validity of Factor B of the ESPQ since the instrument was first published in 1966 (See Coan and Cattell, 1972). The present study was undertaken to determine (1) if there is any similarity in how the Factor B raw scores of the ESPQ measure general mental ability and how the raw scores of the Wechsler Intelligence Scale for Children-Revised(WISC-R) measure general mental ability, and (2) if subjects designated as mentally retarded, learning disabled, or normal differ in their ESPQ Factor B raw scores.

Statement of the Problem

The problem of the study was to assess the validity of

Factor B of the ESPQ in two ways. First, was to determine if there are significant statistical relationships at the .001 level of significance between the ESPQ Factor B raw scores of developmentally disabled subjects and their raw scores on the WISC-R. Second, was to determine if there are significant differences at the .001 level of significance among the ESPQ Factor B raw scores of subjects designated as mentally retarded, learning disabled, or normal. That is, in the second case, the problem was to determine if there are significant differences at the .001 level between the mean ESPQ Factor B raw scores of (1) the mentally retarded subjects and the learning disabled subjects, (2) the mentally retarded subjects and the normal subjects, and (3) the learning disabled subjects and the normal subjects.

Need for the Study

This study fulfills a need in that it attempts to obtain information about Factor B of the ESPQ which might enable a user of the test to better interpret a child's Factor B test score. Inferences about a child's mental ability could greatly influence the child's ultimate welfare. Thus, it is important to know what trust can safely be put on Factor B as a valid measure of mental ability if a test user is to best serve the needs of the child.

Delimitations

The study was delimited to (1) the Joplin Regional Center for Developmentally Disabled, Joplin, Missouri, (2) various schools in Crawford County, Kansas, and (3) Carl Junction

Elementary School, Carl Junction, Missouri. Testing was conducted during the period from February, 1975 through May, 1976. Developmentally disabled and normal children between the ages of 6 years, 0 months and 8 years, 7 months were employed as subjects. Selection of the subjects was delimited to those children who were made available to the investigator and who were free from classes and other obligations during the daily time periods available to the investigator.

Limitations

The study was limited in the following ways:

(1) The results of the study may not generalize to developmentally disabled or normal children in geographical locations other than those chosen for the investigation. Specifically, the raw scores of other developmentally disabled or normal children on Factor B of the ESPQ or on the WISC-R may differ from the raw scores of those subjects employed in the study.

(2) The diagnoses of the subjects as being mentally retarded, learning disabled, or normal were made on a variety of both objective and subjective criteria by different examiners or raters. It was assumed that all independent examiners or raters had sufficient training and experience so as not to make their judgmental decisions significantly different from one another. It is not known what effect this had on the results.

(3) The subject samples employed in the study did not have equal numbers of males and females.

(4) The order of administering the WISC-R and the ESPQ was not balanced. The administration of the ESPQ always followed the administration of the WISC-R.

(5) The ESPQ was individually administered to the developmentally disabled subjects but was administered to the normal subjects in a group situation.

(6) The time and place of testing the subjects was not always the same.

(7) Larger samples and random sampling would have been preferred for the purpose of making the subject samples more representative of their respective populations.

Hypotheses

The four null hypotheses to be tested are as follows:

(1) There is no significant relationship between the ESPQ Factor B raw scores of the developmentally disabled subjects and their raw scores of the WISC-R.

(2) There is no significant difference between the mean ESPQ Factor B raw scores of the mentally retarded subjects and the learning disabled subjects.

(3) There is no significant difference between the mean ESPQ Factor B raw scores of the mentally retarded subjects and the normal subjects.

(4) There is no significant difference between the mean ESPQ Factor B raw scores of the learning disabled subjects and the normal subjects.

The .001 level of significance was adopted as the standard for rejection of all null hypotheses.

Definition of Terms

Developmentally Disabled(DD) Subjects. The term "developmental disability" (Ross, 1975) refers to a disability which is attributable to mental retardation or "to any other condition of a person found to be closely related to mental retardation because such condition results in similar impairment of general intellectual functioning or adaptive behavior to that of mentally retarded persons or requires treatment and services similar to those required for such persons."

Accordingly, the term "developmentally disabled" refers to those subjects in the study who had been given the WISC-R and were either diagnosed as mentally retarded or learning disabled.

Mentally Retarded(MR) Subjects. The term "mentally retarded" refers to those subjects in the study whose Full-Scale IQ on the WISC-R was between 52 and 83. This range coincides with the DSM-II diagnostic classifications of both mild and borderline mental retardation (American Psychiatric Association, DSM-II, 1968, p. 14).

Learning Disabled(LD) Subjects. The term "learning disabled" generally refers to those individuals who manifest a disorder in one or more of the basic psychological processes related to the understanding or use of the spoken or written language. These may be disorders of listening, thinking, writing, spelling, or arithmetic. Problems related to visual, hearing, or motor handicaps, to mental retardation, to behavioral disorder, or to environmental disadvantage are generally

excluded (Lerner, 1976, pp. 9-10).

For the purposes of this study, a diagnosis of a subject as being learning disabled was accepted if the diagnosis was made by a qualified psychologist or mental examiner in the states of Missouri or Kansas.

Normal(NM) Subjects. The term "normal" refers to those subjects in the study who (1) had an IQ between 84 and 122 on the Otis-Lennon Mental Ability Test, Elementary I Level, Form J, (2) were enrolled in regular classes and (3) had not been diagnosed as learning disabled or retarded.

Stens. Cattell (1965, p. 374) defined stens as "units in a standard sten scale in which ten score points are used to cover the population range in fixed and standard deviation intervals, extending from 2½ standard deviations above the mean (sten 10). The span of a sten is ½ sigma and the mean is fixed a 5.5."

WISC-R Scaled Scores. Scales scores on the WISC-R range from 1 to 19 and have a mean of 10 and a standard deviation of 3. According to Wechsler (1974, p. 21) the scaled scores were derived "by preparing a cumulative frequency distribution of raw scores for each age group, normalizing the distribution, and computing the appropriate scaled score for each raw score."

Criterion-Related Validity. According to the American Psychological Association's Standards for Educational and Psychological Tests and Manuals (Jackson & Messick, 1967, p. 176), criterion-related validity refers to the kind of

information gathered when the aim of the test user is to forecast an individual's future standing or to estimate an individual's present standing on some variable of particular significance that is different from the test.

Construct Validity. Construct validity refers to the kind of information gathered when the aim of the test user is to infer the degree to which the individual possesses some hypothetical trait or quality (construct) presumed to be reflected in the test performance (p. 176).

Validity Coefficient. A validity coefficient refers to a correlation between a test and a criterion (Cronbach, 1960, p. 115).

CHAPTER II

REVIEW OF THE LITERATURE

The research relevant to evaluating the validity of Factor B of the ESPQ in this study may be divided into two main areas of investigation: (1) research related to demonstrating the validity of Factor B of the ESPQ as a measure of general mental ability, and (2) research related to demonstrating the validity of the WISC-R as a measure of general mental ability.

Research Related to Demonstrating the Validity of Factor B of the ESPQ

It has been demonstrated in at least two criterion-related validity studies that Factor B of the ESPQ had some validity for assessing children's current status on a number of variables other than general mental ability. Dielman, Cattell, and Lepper (1971) found significant negative correlations between Factor B of the ESPQ problem behaviors as speech problems, social withdrawal, paranoid tendencies, neurasthenia, and disciplinary problems. Jackson (1972) administered the ESPQ and the Stanford Reading Achievement Test to 325 normal second graders and found that Factor B was the only factor on the instrument which significantly discriminated between low achievers(low Factor B score) and high achievers(high Factor B score).

No studies have yet been reported in the literature

which investigated the relationship between Factor B of the ESPQ and a standardized test of general intelligence. However, one study has related Factor B of the ESPQ to Factor B of the High School Personality Questionnaire, a better established personality test for children in the 12-18 year age range. Coan and Cattell (1959) administered the ESPQ and High School Personality Questionnaire to a sample of 92 children between the ages of 8 years, 0 months and 10 years, 11 months. They reported a significant product-moment correlation of .32 (level of significance was unreported) between Factor B of the ESPQ and Factor B of the HSPQ.

Other studies (Karson & Pool, 1957; Tamkin, 1967; Lessing & Zagorin, 1959; and Fleishman & Fine, 1971) have investigated the correlations between various intelligence tests and Factor B as it appears on other personality questionnaires for older children and adults published by the Institute of Personality and Ability Testing(IPAT), Champaign, Illinois. Substantial positive correlations between Factor B and the intelligence tests were generally obtained. None of these studies employed Factor B of the ESPQ. Thus, the studies offer information only of a suggestive and tangential nature concerning the validity of Factor B of the ESPQ since, except for the one study above (Coan & Cattell, 1959), the correlations between Factor B of the ESPQ and Factor B as it appears on other IPAT personality questionnaires is unknown.

Finally, the ESPQ manual (Coan & Cattell, 1972, p. 23) reported a direct concept validity coefficient of .78 for

Factor B. The concept validity coefficient was obtained by correlating the actual B scale value with the pure factor as determined by factor analysis. Coan and Cattell (p. 21) also reported reliability coefficients of .68 for Factor B which were obtained with the separate samples when estimated by the split-half method and corrected by the Spearman-Brown formula for full-test length.

Research Related to Demonstrating the Validity of the WISC-R

Because the WISC-R is used as a criterion of general mental ability in this study, it is important to review some of the literature which give support for the instrument as a suitable criterion.

The WISC-R is a revision of the 1949 WISC. The WISC has gained wide acceptance as a reliable and valid measure of general mental ability (Littell, 1960; and Zimmerman & Woo-Sam, 1972). Despite a great similarity between the WISC and WISC-R, the WISC-R is still a relatively new test and has not had as lengthy a research history as the WISC.

The WISC-R involved the addition of new items to strengthen the reliability of each test and according to Wechsler (1974, p. 10) "the modification or elimination of items felt by some test users to be ambiguous, obsolete, or differentially unfair to particular groups of children."[1]

For the Verbal IQ, Wechsler (p. 28) reported reliability

---

[1]See Wechsler (1974, pp. 10-16) for a more detailed description of changes in item content, administration and scoring.

coefficients of .91, .92, and .92 for the 6½, 7½, and 8½ year age groups, respectively. For the Performance IQ, he reported reliability coefficients of .91, .90, and .91 for the same respective age groups. For the Full Scale IQ, he reported reliability coefficients of .95 for each of the 6½, 7½, and 8½ year age groups. The average reliability coefficient of the Verbal IQ, Performance IQ, and Full Scale IQ of the WISC-R for all age groups from 6½ to 16½ years was .94, .90, and .96, respectively.

Stability coefficients for the WISC-R were obtained by retesting a total of 303 children in the standardization sample and correlating initial scores with final scores. The obtained stability coefficients were very similar in magnitude to the reliability coefficients and provide further evidence of the reliability of the WISC-R (pp. 30-31).

Wechsler (p. 49) reported correlations of the WISC-R with both the WPPSI and the Stanford-Binet Intelligence Scale for children six years of age. When the WISC-R was correlated with the WPPSI, the verbal subtests of the WISC-R were found to correlate better with the verbal subtests of the WPPSI and the performance subtests of the WISC-R were found to correlate better with the performance subtests of the WPPSI. The correlation of .78 between the WISC-R Performance IQ and the WPPSI Full Scale IQ was higher than the correlation of .73 between the WISC-R Verbal IQ and the WPPSI Full Scale IQ. The correlation between the WISC-R Full Scale IQ and the WPPSI Full Scale IQ was .82.

When the WISC-R was correlated with the Stanford-Binet, the WISC-R Vocabulary subtest had the highest of the subtest correlations with the Stanford-Binet (.75), followed by Picture Arrangement (.74), Mazes (.69), and Comprehension (.66). The correlation coefficients of the WISC-R Verbal, Performance, and Full Scale IQs with the Stanford-Binet IQ were .77, .74, and .82, respectively (p. 52).

Some researchers have investigated the validity of the WISC-R by employing factor analysis. Kaufman (1974) factor analyzed the WISC-R with the standardization sample of 200 boys and girls in each of the age groups from 6½ to 16½ years. He reported its structure to be very similar to that of the 1949 WISC. He concluded (p. 147) that "the structure of the WISC-R is at the same time both consistent with, and a decided improvement over, the structure of its highly successful predecessor" (the WISC).

In a related study, Kaufman and Hagen (1975) factor analyzed the WISC-R with a group of 80 retarded children and found essentially the same results Kaufman (1974) found with the normal children in the standardization sample.

In another factor analytical study, Wallbrown, Blaha, Wallbrown, and Engin (1975) factor analyzed the WISC-R with the standardization sample across all age groups. Their findings were, to a large degree, very similar to Kaufman's (1974) findings. Walbrown, et. al., concluded that their findings "provide a substantial degree of construct validity for the WISC-R as a measure of general intelligence." (p. 233)

Summarily, the literature to date gives support for the WISC-R as a reliable and valid measure of general mental ability. Evidence in the literature for the reliability and validity of Factor B of the ESPQ is less substantial.

CHAPTER III

RESEARCH DESIGN

Test Instruments

Early School Personality Questionnaire(ESPQ). Including Factor B, the ESPQ consists of thirteen personality scales; each is designed to measure a relatively independent personality dimension (Coan & Cattell, 1972). Form A, employed in this study, is divided into two parts: A_1 and A_2. Each part contains eighty items--eight items on each part for Factor B and six items on each part for each of the other twelve personality scales. The Factor B test items, as they appear on parts A_1 and A_2 of the ESPQ, are presented in Appendix A, p. 38.

The ESPQ is designed for group or individual administration to children between the ages of six years, 0 months to 8 years, 11 months, 15 days. The questions are read aloud to a child and his responses are recorded on a "non-reading" answer form. Each page contains twenty rectangular boxes arranged in two columns. In the middle boxes are item numbers, as well as pictures, which can serve the same purpose of locating a particular items for younger subjects. There is an A at the left end of each box and a B at the right end. During group administration, each child is asked to indicate his response by drawing a line through either the A or the B in each box. During individual administration, the child may mark his own responses or the examiner may mark the child's responses for him.

After the administration of the ESPQ, the raw scores of each of the personality scales, including Factor B, are converted to sten scores and recorded on a standard form which makes provision for a profile comparison of all scores.

Wechsler Intelligence Scale for Children-Revised(WISC-R). The WISC-R consists of the same twelve tests that constituted the 1949 WISC. The six verbal tests are: Information, Similarities, Arithmetic, Vocabular, Comprehension, and Digit Span. The six performance tests are: Picture Completion, Picture Arrangement, Block Design, Object Assembly, Coding, and Mazes. All of the subtests are regularly administered with the exception of Digit Span and Mazes, which are optional. In contradistinction to the WISC, these optional tests neither contribute to nor detract from a subject's final score.

The examiner gives all WISC-R instructions aloud and records the subject's responses on a standard record form. After the administration of the WISC-R, the raw scores for each of the five verbal and five performance tests are converted to scaled scores and are summed to obtain a sum of scaled scores which are designated as a Verbal Score, a Performance Score; and a Full-Scale Score, respectively. The three scaled scores are then converted to Verbal, Performance, and Full Scale IQs, each with a mean of 100 and a standard deviation of 15.

Subjects

The sex, chronological age, IQ, and raw score of each subject employed in the study are given in Appendix B and

Appendix C, pp. 41, 42. According to school records and psychological evaluations, none of the subjects manifested visual or auditory handicaps.

Developmentally Disabled(DD) Subjects. A summary of the characteristics of the DD subjects is presented in Table I, p. 18. It can be observed that six of the subjects were female and 34 were male. The mean chronological age was 7 years, 3 months and the ages ranged from 6 years, 0 months to 8 years, 7 months. The mean ESPQ Factor B raw score was 9.62 and the scores ranged from four to 15. The mean WISC-R Full-Scale was 84.50 and the IQs ranged from 57 to 123.

The first 30 of the 40 DD subjects listed in Appendix B, p. 40 were obtained from the Joplin Regional Center for Developmentally Disabled, Joplin, Missouri. The subjects were seen at various times and tested at the center over a period of a year. It was estimated from Raum (1976, p. 41) that there are approximately 800 developmentally disabled children between the ages of six and nine in the area served by the center. After the subjects were given a psychological evaluation at the center, they were diagnosed either as mentally retarded or as learning disabled.

The last ten subjects listed in Appendix B, p. 40 were obtained from a population of about 92 mentally retarded and 155 learning disabled children in the area of Crawford County, Kansas. The children had been diagnosed as either mentally retarded or as learning disabled by school psychologists and were enrolled in classes of various schools located in the

county. Six of the subjects tested in Crawford County were not included in the sample because the WISC-R raw scores of the subjects were not available.

Of the 40 DD subjects, 15 were diagnosed as mentally retarded and 25 were diagnosed as learning disabled.

Mentally Retarded(MR) Subjects. It can be observed in Table I, p. 18 that two of the MR subjects were female and 18 were male. The mean chronological age was 7 years, 3 months and the ages ranged from 6 years, 0 months to 8 years, 7 months. The mean ESPQ Factor B raw score was 7.05 and the scores ranged from four to nine. The mean WISC-R Full-Scale IQ was 67.2 and the IQs ranged from 53 to 83.

The subject characteristics for each MR subject were presented in Appendix C, p. 42. It can be observed in Appendix C that the first 15 MR subjects listed are the same MR subjects in the DD group. The last five additional MR subjects were obtained from various schools located in Crawford County, Kansas, from a population of about 92 MR children.

Learning Disabled(LD) Subjects. It can be observed in Table I that four of the LD subjects were female and 22 were male. The mean chronological age was 7 years, 2 months and the ages ranged from 6 years, 1 month to 8 years, 7 months. The mean ESPQ Factor B raw score was 11.19 and the scores ranged from eight to 15. The mean WISC-R Full-Scale IQ was 95.4 and the IQs ranged from 79 to 123.

In Appendix C, p. 42 it can be observed that the first

25 of the LD subjects listed are the same subjects as in the DD group. The last subject was obtained, as above, from a population of about 155 learning disabled children in Crawford County, Kansas.

TABLE I

SUMMARY OF
SUBJECT CHARACTERISTICS

S	N	M	F	CA Mean	CA Range	B Mean	B Range	IQ Mean	IQ Range
DD	40	34	6	7-3	6-0 to 8-7	9.62	4-15	84.5*	53-123*
MR	20	18	2	7-3	6-0 to 8-7	7.05	4-9	67.2*	53-83*
LD	26	22	4	7-2	6-1 to 8-7	11.19	8-15	95.4*	79-123*
NM	16	16	0	7-3	6-9 to 8-0	11.75	9-15	105.0**	91-122**

*WISC-R Full-Scale IQ
**Otis-Lennon IQ
S = Subjects
N = Number of subjects
M = Male
F = Female
CA = Chronological age in years and months
B = ESPQ Factor B raw score

Normal (NM) Subjects. It can be observed in Table I that all of the NM subjects were male. The mean chronological age was 7 years, 3 months and the ages ranged from 6 years, 9 months to 8 years, 0 months. The mean ESPQ Factor B raw score was 11.75 and the scores ranged from nine to 15. The mean Otis-Lennon IQ was 105 and the IQs ranged from 91 to 122.

The 16 NM subjects listed in Appendix C, p. 42, were obtained from a population of about 300 first and second grade

students enrolled in regular classes at Carl Junction Elementary School, Carl Junction, Missouri. Six were first graders and ten were second graders.

Procedure

The DD, LD, MR, and NM subject groups were formed on the basis of aforementioned criteria (See Definition of Terms, pp. 5-7). WISC-R test scores and diagnoses were available on the DD, LD, and MR subjects and Otis-Lennon scores were available on the NM subjects. The WISC-R had been administered by the investigator to all of the subjects tested at the Joplin Regional Center and by school psychologists to all of the subjects tested in Crawford County, Kansas. The Otis-Lennon had been administered to the NM subjects by the Carl Junction Elementary School counselor and other school personnel. All of the NM subjects were enrolled in regular classes. In the judgment of the elementary school counselor, neither the school records nor the classroom performance of the NM subjects indicated that any of them were manifesting learning difficulties in the classroom. None of the NM subjects had been diagnosed as learning disabled or mentally retarded.

Form A of the ESPQ was administered by the investigator to all of the subjects. The ESPQ was administered to the NM subjects in a group situation. All other subjects were individually administered the ESPQ as each became available for testing. The administration of the ESPQ always followed the administration of the WISC-R or the Otis-Lennon. The time interval between the administration of the WISC-R and the

ESPQ averaged one week in the case of the subjects tested at the Joplin Regional Center and averaged five months in the case of the subjects tested in Crawford County, Kansas. The time interval between the administration of the Otis-Lennon and the ESPQ in the case of the NM subjects was six months.

Statistical Analysis. In order to obtain the validity coefficients for Factor B of the ESPQ, product-moment correlations were calculated between the Factor B raw scores of the 40 DD subjects listed in Appendix B, p. 40, and their raw scores on the various tests of the WISC-R. Specifically, the following product-moment correlations were calculated:

(1) between the ESPQ Factor B raw scores and the raw scores of each separate subtest of the WISC-R, excluding the Mazes subtest;

(2) between the ESPQ Factor B raw scores and the raw score sum of all Verbal subtests of the WISC-R;

(3) between the ESPQ Factor B raw scores and the raw score sum of all Performance subtests of the WISC-R; and

(4) between the ESPQ Factor B raw scores and the sum of raw scores over all of the WISC-R subtests (Verbal subtest raw scores + Performance subtest raw scores).

The product-moment correlation coefficients were computed using the formula (Guilford, 1965, p. 97):

$$r = \frac{N\Sigma XY - \Sigma X \Sigma Y}{\sqrt{(N\Sigma X^2 - (\Sigma X)^2)(N\Sigma Y^2 - (\Sigma Y)^2)}}$$

where r = product-moment correlation coefficient

In order to test the significance of each obtained coefficient, t tests were conducted on each obtained product-moment correlation coefficient. The criterion of acceptance was set at the .001 level of confidence. The t tests were computed using the formula (Guilford, 1965):

$$t = r\sqrt{\frac{N - 2}{1 - r^2}}$$

where df = N - 2

In order to test whether there were significant differences among the mean ESPQ Factor B raw scores of the mentally retarded, learning disabled, or normal subjects, the Kruskal-Wallis one-way analysis of variance rank sums test was employed. The Kruskal-Wallis test was chosen in order to avoid making any assumptions about the normality of the distributions of the respective populations (See Limitations section).

With the criterion of acceptance set at the .001 level of confidence, the Kruskal-Wallis H statistic was computed using the following formula, corrected for ties (Siegel, 1956, p. 188):

$$H = \frac{\frac{12}{N(N + 1)} \sum_{j=1}^{k} \frac{R_j}{n_j} - 3(N + 1)}{1 - \frac{\Sigma T}{N^3 - N}}$$

where $T = t^3 - t$ (when t is the number of ties observations in a tied group of scores

N = the total number of observations in all of the k

samples together

k = number of subject samples

n_j = number of cases in the jth subject sample

R_j = sum of ranks in the jth subject sample (column)

df = k - 1

Although the Kruskal-Wallis procedure tests whether treatments are equivalent or not, it does not test which treatments significantly differ from one another (Hollander & Wolfe, 1973, p. 125). Thus, the distribution-free multiple comparisons procedure was employed in order to determine if there were significant differences between the mean ESPQ Factor B raw scores of (1) the mentally retarded subjects and the learning disabled subjects, (2) the mentally retarded subjects and the normal subjects, and (3) the learning disabled subjects and the normal subjects. The multiple comparisons test, a variation of the Kruskal-Wallis test, is also a nonparametric test and foregoes any assumption that the underlying populations are normally distributed.

Hollander and Wolfe (1973, p. 125) gave the following procedure for the distribution-free multiple comparisons test:

With a .001 probability error rate ($\alpha = .001$),

to decide $r_u \lessgtr r_v$ if

$$\left|\bar{R}_u - \bar{R}_v\right| \gtrless z_{(\alpha/[k(k-1)])} \left[\frac{N(N+1)}{12}\right]^{\frac{1}{2}} \left(\frac{1}{n_u} + \frac{1}{n_v}\right)^{\frac{1}{2}}$$

where r = unknown treatment effect

$\bar{R}_u, \bar{R}_v$ = mean ranks corresponding to each of the (u,v) pairs

of subjects

α = probability error rate

k = number of subject samples

N = total number of observations in all of the k subject samples

n_u, n_v = number of observations in each subject sample

CHAPTER IV

ANALYSIS OF DATA

Results

Validity Coefficients. The validity coefficients for factor B of the ESPQ are presented in Table II, p. 25. It can be observed that the highest validity coefficient (.80) for Factor B was obtained from the correlation of the Factor B raw scores with the Sum of Verbal Test raw scores of the WISC-R. The next highest validity coefficient for Factor B was obtained from the correlation of the Factor B raw scores with the raw scores of the Vocabulary subtest of the WISC-R (.78). Next was the validity coefficient of .75 obtained from the correlation of the Factor B raw scores with the raw score sum of all the WISC-R subtests (Verbal + Performance Tests).

The results of the t tests performed indicated that all of the validity coefficients of Factor B were significant at the .001 level with the exception of the validity coefficient (.30) obtained from the correlation of the Factor B raw scores with the raw scores of the WISC-R Coding subtest.

Kruskal-Wallis Rank Sums Test and Multiple Comparisons Test. The value of the Kruskal-Wallis H statistic, corrected for ties, was 225.59 (See Appendix D, pp. 44, 45). Reference to Table C in Siegel (1956, p. 249) indicated that the probability of an H as large as 225.59, with two degrees of freedom, is equal to or less than .001. Thus, at the .001 level

TABLE II

VALIDITY COEFFICIENTS OF FACTOR B OF THE ESPQ OBTAINED BY CORRELATING FACTOR B RAW SCORES WITH THE RAW SCORES OF THE WISC-R
N = 40

	Factor B Raw Score			
WISC-R Test*	Validity Coefficient	t	df**	Level of Significance
Information	.56	4.17	38	.001
Similarities	.53	3.85	38	.001
Arithmetic	.63	5.00	38	.001
Vocabulary	.78	7.68	38	.001
Comprehension	.63	5.00	38	.001
Digit Span	.63	5.00	37	.001
Picture Completion	.54	3.96	38	.001
Picture Arrangement	.62	4.87	38	.001
Block Design	.64	5.14	38	.001
Object Assembly	.56	4.17	38	.001
Coding	.30	1.94	38	Not Significant
Sum of Verbal Tests	.80	8.22	38	.001
Sum of Performance Tests	.62	4.87	38	.001
Verbal + Performance Tests	.75	6.99	38	.001

*WISC-R raw scores
**df = degrees of freedom

TABLE III

RESULTS OF DISTRIBUTION-FREE MULTIPLE COMPARISONS THAT

Subjects Compared	$\left\|\bar{R}_u - \bar{R}_v\right\|$	$z(\alpha/[k(k-1)])\left[\frac{N(N+1)}{12}\right]^{\frac{1}{2}}\left(\frac{1}{n_u}+\frac{1}{n_v}\right)^{\frac{1}{2}}$	Decision
MR-LD	27.188	18.728	$r_u \neq r_v$ *
MR-NM	30.994	21.121	$r_u \neq r_v$ *
LD-NM	3.806	20.008	$r_u = r_v$ **

* Significant where probability error rate = .001
** Not significant
MR = Mentally retarded
LD = Learning disabled
NM = Normal
$\bar{R}_u, \bar{R}_v$ = Mean ranks corresponding to each of the (u,v) pairs of subjects
α = Probability error rate
k = Number of subject samples
n_u, n_v = Number of observations in each subject sample
r_u, r_v = Unknown treatment effects corresponding to each of the (u,v) subject sample pairs
Decision = decide $r_u \neq r_v$ if $\left|\bar{R}_u - \bar{R}_v\right| \geq z(\alpha/[k(k-1)])\left[\frac{N(N+1)}{12}\right]^{\frac{1}{2}}\left(\frac{1}{n_u}+\frac{1}{n_v}\right)^{\frac{1}{2}}$

of confidence, the MR, LD, and NM subjects significantly differ in their mean ESPQ Factor B raw scores.

The results of the application of the distribution-free multiple comparisons test are presented in Table III, p. 26 (See also Appendix E, p. 48). According to the multiple comparisons equation the mean Factor B raw scores corresponding to each of the (u,v) subject pairs are significantly different at a probability error rate[1] of .001 if the absolute differences of the average ranks of the ESPQ Factor B raw scores $\left|\bar{R}_u - \bar{R}_v\right| \geq$

$$z_{(\alpha/[k(k-1)])}\left[\frac{N(N+1)}{12}\right]^{\frac{1}{2}}\left(\frac{1}{n_u}+\frac{1}{n_v}\right)^{\frac{1}{2}}$$

Thus, for the MR-LD subject comparison, it can be observed in Table III that this condition of the multiple comparisons equation was satisfied where the absolute value of 27.188 was greater than the value of 18.728. Hence, it was decided that the treatment effect $r_u \neq r_v$, or, that the mean ESPQ Factor B raw scores of the LD subjects was significantly higher than the mean ESPQ Factor B raw scores of the MR subjects when the probability error rate was .001.

---

[1]Note: The term "probability error rate", used in this statistic, closely corresponds to the term "level of confidence" used with parametric statistical tests. Since the multiple comparisons procedure is a nonparametric test, this change in statistical terminology becomes necessary (See Hollander and Wolfe (1973, p. 448) for a technical definition of "probability error rate".).

For the MR-NM subject comparison, the absolute value of 30.994 was greater than the value of 21.121. Thus, this met the condition of the multiple comparisons equation to decide that the treatment effect $r_u \neq r_v$, or, that the mean ESPQ Factor B raw scores of the NM subjects was significantly higher than the mean ESPQ Factor B raw scores of the MR subjects when the probability error rate was .001.

For the LD-NM subject comparison, it can be observed in Table III that with a probability error rate of .001, the absolute value of 3.806 was less than the value of 20.008. Hence, this met the condition of the multiple comparisons equation to decide that the treatment effect $r_u = r_v$, or, that the mean ESPQ Factor B raw scores of the LD and NM subjects were not significantly different when the probability error rate was .001.

CHAPTER V

DISCUSSION

Conclusions and Summary

The specific concern of this study was to investigate the validity of Factor B of the ESPQ as a measure of general mental ability (1) by obtaining validity coefficients for Factor B, and (2) by determining if there were significant differences among the mean Factor B raw scores of subjects designated as mentally retarded, learning disabled, or normal. In the first case, the determination was made by computing the product-moment correlation coefficients between the ESPQ Factor B raw scores of 40 developmentally disabled subjects and their raw scores on the WISC-R. In the second case, the determination was made by employing the Kruskal-Wallis rank sums test and the distribution-free multiple comparisons test to determine if there were significant differences between the mean ESPQ Factor B raw scores of (1) the mentally retarded subjects and the learning disabled subjects, (2) the mentally retarded subjects and the normal subjects, and (3) the learning disabled subjects and the normal subjects. The criterion of acceptance was set at the .001 level of confidence in all instances.

All of the validity coefficients for Factor B of the ESPQ were found to be significant at the .001 level with the exception of the validity coefficient obtained from the

correlation of the ESPQ Factor B raw scores with the raw scores of the WISC-R Coding subtest. Accordingly, the null hypothesis was rejected, it was concluded that the ESPQ Factor B raw scores, to a degree, measured the general mental ability of the developmentally disabled subjects much like the raw scores of the WISC-R measured the general mental ability of the developmentally disabled subjects in the study.

In order to determine which subject samples significantly differed in mean ESPQ Factor B raw scores from one another, the distribution-free multiple comparisons test was employed. For the MR-LD and MR-NM subject sample comparisons, the absolute differences of the average ranks $\left|\bar{R}_u - \bar{R}_v\right|$ was greater than the value $z_{(\alpha/\left[k(k-1)\right])}\left[\frac{N(N+1)}{12}\right]^{\frac{1}{2}}\left(\frac{1}{n_u}+\frac{1}{n_v}\right)^{\frac{1}{2}}$.

With the probability error rate set at .001, this satisfied the conditions of the multiple comparisons equation to decide that the treatment effects r_u, r_v were not equal. Thus, the mean ESPQ Factor B raw scores of the learning disabled and normal subjects were significantly higher than the mean ESPQ Factor B raw scores of the mentally retarded subjects. Accordingly, it can be concluded that the learning disabled subjects and normal subjects were higher than the mentally retarded subjects in mean level of general mental ability as measured by the ESPQ Factor B raw scores.

For the LD-NM subject sample comparison, the absolute differences of the average ranks $\left|\bar{R}_u - \bar{R}_v\right|$ was less than the value

$$z_{(\alpha/\left[k(k-1)\right])}\left[\frac{N(N+1)}{12}\right]^{\frac{1}{2}}\left(\frac{1}{n_u}+\frac{1}{n_v}\right)^{\frac{1}{2}}.$$

With the probability error rate set at .001, this satisfied the condition of the multiple comparisons equation to decide that the treatment effects r_u, r_v were equal. Hence, the mean ESPQ Factor B raw scores of the learning disabled subjects and normal subjects did not significantly differ. Accordingly, it can be concluded that the learning disabled subjects and normal subjects were about equivalent in mean level of general mental ability as measured by the ESPQ Factor B raw scores.

The latter finding is not completely unexpected since learning disabled children are not generally differentiated from normal children on the basis of general mental ability, but on the basis of other criteria such as specific disorders in basic psychological processes related to listening, thinking, writing, spelling, or arithmetic (Lerner, 1976, pp. 9-10). Thus, it can be concluded that the ESPQ Factor B raw scores apparently were not an accurate measure of those specific aspects of the psychological processes which differentiated the learning disabled child from the normal child in the study.

In summary, the general findings suggest that the ESPQ Factor B raw scores, to a good degree, were a valid measure of the general mental ability of the subjects employed in the study since (1) the validity coefficients obtained for Factor B were generally substantial with the sample of developmentally disabled subjects employed, and (2) significant differences in mean ESPQ Factor B raw scores were measured among the subjects designated as mentally retarded,

learning disabled, or normal.

Recommendations for Further Research

In order to make more valid statistical inferences to the respective populations of the subject samples employed in the study, this investigator recommends that random samples be employed in a replication of the study. Additionally, other studies might employ larger, stratified subject samples in order to insure that representative proportions of relevant variables are included in the study. Some relevant variables to include in the stratified samples might be age, grade placement, sex, race, and geographical region.

This investigator further recommends that studies be conducted to investigate the validity of the other personality factors on the ESPQ by determining the relationships between the ESPQ factors and other variables such as (1) the presence of behavioral disorders, learning disabilities, or mental retardation, (2) the prediction of academic achievement, or (3) other personality measures.

BIBLIOGRAPHY

BIBLIOGRAPHY

BOOKS

American Psychiatric Association. Diagnostic and statistical manual of mental disorders(DSM-II). Washington, D. D.: American Psychiatric Association, 1968.

American Psychological Association. Standards for educational and psychological tests and manuals. In D. Jackson & Messick (Eds.), Problems in Human Assessment. New York: McGraw-Hill, 1967.

Cattell, R. B. The scientific analysis of personality. Baltimore: Penquin Books, 1965.

Coan, R. W., & Cattell, R. B. Manual for the Early School Personality Questionnaire. Champaign, Illinois: Institute for Personality and Ability Testing, 1972.

Cronbach, L. J. Essentials of Psychological testing (2nd ed.). New York: Harper & Brothers, 1960.

Guilford, J. P. Fundamental statistics in psychology and education. New York: McGraw-Hill, 1965.

Hollander, M., & Wolfe, D. A. Nonparametric statistical methods. New York: John Wiley & Sons, 1973.

Lerner, J. W. Children with learning disabilities (2nd ed.). Boston: Houghton Mifflin Co., 1976.

Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.

Wechsler, D. Manual for the Wechsler Intelligence Scale for Children-Revised. New York: Psychological Corporation, 1974.

PERIODICALS

Coan, R. W., & Cattell, R. B. The development of the Early School Personality Questionnaire. Journal of Experimental Education, 28, 2,1959, 143-152.

Dielman, T. E., Cattell, R. B., & Lepper, C. Personality correlates of behavior problems in early childhood. Personality: An International Journal, 2, 2, 1971, 141-147.

Fleishman, J. J., & Fine, B. J. Note on cognitive factors related to Factor B on the 16PF Test. Psychological Reports, 1971, 29, 1075-1077.

Jackson, R. N. A study of six personality factors in reading achievement. Ph.D. Dissertation, The Ohio State University, 1972. (ERIC Document Reproduction Service No. ED 065 853)

Karson, S., & Pool, K. B. The construct validity of the Sixteen Personality Factors Test. Journal of Clinical Psychology, 1957, 13, 245-252.

Kaufman, A. S. Factor analysis of the WISC-R at 11 age levels between 6½ and 16½ years. Journal of Consulting and Clinical Psychology, 43, 2, 135-142.

Kaufman, A. S., & Hagen, J. V. Factor analysis of the WISC-R for a group of mentally retarded children and adolescents. Journal of Consulting and Clinical Psychology, 43, 5, 1975, 661-667.

Lessing, E. E., & Zargorin, S. W. Correlation between the Lorge-Thorndike IQ and Factor B of the IPAT Childrens

Personality Questionnaire. Psychological Reports, 1969, 569-570.

Littell, W. M. The Wechsler Intelligence Scale for Children: Review of a decade of research. Psychological Bulletin, 1968, 57, 132-156.

Tamkin, A. S. Correlation between IQ and Factor B of the Sixteen Personality Factor Test with psychiatric patients. Journal of Clinical Psychology, 1967, 23, 486-488.

Wallbrown, F. H., Blaha, J., Wallbrown, J. O., & Engin, A. W. The hierarchial factor structure of the Wechsler Intelligence Scale for Children-Revised. The Journal of Psychology, 1975, 89, 223-235.

Zimmerman, I. L., & Woo-Sam, J. Research with the Wechsler Intelligence Scale for Children: 1960-1970. Psychology in the Schools, 1972, 232-271 (Special Monograph Supplement)

OTHER PUBLICATIONS

Ross, E. C. Word on Washington. Washington, D. C., United Cerebral Palsy Association, 6, 10, 1975.

UNPUBLISHED MANUSCRIPTS

Raum, Jim D. Region Five Plan for the Developmental Disabilities: 1977. Title I, Part C, P. L. 94-103, 1976.

APPENDIX A

FACTOR B INTELLIGENCE ITEMS ON PARTS A_1 AND A_2 OF FORM A OF THE ESPQ*

Part A_1	Part A_2
Is a large house : (A) a big house, or (B) a small house?	Is a pretty picture : (A) beautiful, or (B) ugly?
Is a butterfly : (A) a bird, or (B) an insect?	Is a daisy : (A) a tree, or (B) a flower?
What does a house always have : (A) a chimney, or (B) a roof?	What does a car always have : (A) a radio, or (B) an engine?
Jenny is smarter than Louise. Louise is smarter than Rose. Who is smarter : (A) Jenny, or (B) Rose?	Jane is older than Helen. Helen is older than Alice. Who is older: (A) Jane, or (B) Alice?
If something is true, is it : (A) correct, or (B) false?	Is a rapid horse : (A) a fast horse, or (B) a slow horse?
Is a giraffe : (A) a jungle animal, or (B) a farm animal?	Is satin : (A) cloth, or (B) paper?
What does shoes always have : (A) shoestrings, or (B) soles?	Which of these things are clothing : (A) glasses or (B) trousers?
If Mary is my father's daughter, is Mary : (A) my mother, or (B) my sister?	Harry is taller than John. John is taller than Bill. Who is shorter : (A) Harry, or (B) Bill?

* From Form A of the Early School Personality Questionnaire (Coan & Cattell, 1972)

APPENDIX B

SUBJECT CHARACTERISTICS AND RAW SCORES OF DEVELOPMENTALLY DISABLED(DD) SUBJECTS

Sub. No.	Sex	CA	B	IQ	WISC-R Raw Scores													
					VS	PS	VS + PS	I	S	A	V	Cm	DS	PC	PA	BD	OA	Cd
1	f	6-6	12	93	49	46	95	5	9	4	18	13	10	11	5	6	8	16
2	f	6-10	11	89	48	48	96	7	10	6	17	8	8	7	5	4	6	26
3	m	6-10	8	85	39	71	110	4	6	5	16	8	4	3	6	7	10	45
4	m	8-0	13	100	76	88	164	10	10	7	34	15	6	14	23	16	13	19
5	m	6-1	13	98	58	24	82	8	8	5	25	12	2	8	2	0	2	12
6	m	6-9	12	95	56	58	114	7	9	6	23	11	6	8	6	5	10	29
7	f	6-8	10	85	30	82	112	5	2	6	13	4	6	8	7	14	8	45
8	m	6-10	11	98	51	79	130	8	8	5	22	8	6	9	10	10	14	36
9	m	6-1	10	82	35	19	54	4	7	1	17	6	2	6	2	1	2	8
10	m	6-10	12	98	39	99	138	5	2	6	13	13	6	11	11	14	16	47
11	m	7-7	15	101	54	108	162	6	10	6	20	12	6	14	22	15	23	34
12	m	7-11	9	79	45	65	110	6	7	4	17	11	8	8	16	6	11	24
13	m	7-9	10	92	54	90	144	7	10	8	18	11	8	16	16	6	8	44
14	m	6-4	9	104	37	87	124	5	6	3	17	6		16	7	15	14	35
15	m	8-2	12	100	46	103	149	12	5	5	18	6	8	7	18	23	22	33
16	m	7-3	9	86	33	81	114	5	1	6	16	5	3	14	6	10	18	33
17	f	7-11	9	57	20	42	62	4	0	2	10	4	3	6	3	4	6	23
18	m	6-11	8	75	28	62	90	3	1	1	15	8	2	12	1	3	14	32
19	m	7-3	9	60	27	25	52	6	0	3	12	6	2	6	2	0	3	14
20	m	6-4	9	81	26	31	57	3	5	3	8	7	1	8	2	0	6	15
21	m	7-2	8	63	17	45	62	2	0	2	11	2	2	8	6	3	7	21
22	m	7-8	8	63	19	56	75	3	2	5	6	3	1	9	2	2	11	32
23	m	6-0	5	69	15	19	34	4	0	2	7	2	2	4	1	2	3	9
24	m	6-10	9	76	26	38	64	4	3	1	14	4	2	7	0	4	12	15
25	m	6-11	5	64	18	47	65	4	0	1	9	4	2	8	2	0	5	32
26	m	7-8	6	73	37	54	91	4	9	1	16	7	6	4	6	4	12	28
27	m	6-8	7	62	22	32	54	4	2	0	9	7	2	5	2	0	3	22
28	m	6-6	4	65	34	28	62	7	6	2	13	6	2	3	1	0	4	20
29	m	6-4	4	59	7	29	36	1	1	0	5	0	0	3	0	0	1	25
30	m	7-8	5	60	20	22	42	4	0	4	11	5	2	5	2	2	7	10
31	m	7-7	14	123	61	121	182	6	10	4	32	9	8	13	29	10	22	47
32	m	6-3	12	123	53	92	145	10	4	5	26	8	8	8	6	15	2	42
33	f	7-5	11	101	54	85	139	5	3	5	22	9	6	7	12	10	22	34
34	m	8-2	12	97	44	104	148	6	8	4	20	6	6	9	11	18	23	43
35	m	8-7	15	88	54	62	116	8	5	5	28	8	5	8	12	11	11	20
36	m	7-7	12	97	45	75	120	6	4	4	26	5	6	7	9	11	14	34
37	f	7-9	11	89	43	64	107	6	6	5	20	6	6	8	18	3	15	20
38	m	6-9	8	85	23	19	42	2	5	4	10	2	4	3	4	0	4	8
39	m	7-4	12	98	48	87	135	4	6	7	19	12	7	10	6	19	20	32
40	m	8-7	8	71	40	57	97	7	7	4	12	10	3	5	5	0	14	33
Mean		7-3	9.62	84.50	38.27	61.10	99.37	5.43	4.92	3.92	16.62	7.22	4.54	8.15	7.60	6.80	11.12	27.42
SD		0-9	2.85	17.03	15.28	28.64	39.88	2.28	3.44	2.04	6.86	3.49	2.59	3.50	6.98	6.41	6.60	11.50
Range		6-0 to 8-7	4-15	53-123	7-76	15-121	34-182	1-12	0-10	0-8	5-34	0-15	0-10	3-16	0-28	0-23	1-23	8-47

CA = Chronological age in years and months
f = Female
m = Male
B = ESPQ Factor B raw score
IQ = WISC-R Full-Scale IQ
VS = Sum of Verbal subtests
PS = Sum of Performance subtests
I = Information
S = Similarities
A = Arithmetic
V = Vocabulary
Cm = Comprehension
DS = Digit Span
PC = Picture Completion
PA = Picture Arrangement
BD = Block Design
OA = Object Assembly
Cd = Coding

APPENDIX C

SUBJECT CHARACTERISTICS AND RAW SCORES OF THE MENTALLY RETARDED, LEARNING DISABLED, AND NORMAL SUBJECTS

MR Subjects					LD Subjects					NM Subjects				
Subject No.	Sex	CA	B	IQ*	Subject No.	Sex	CA	B	IQ*	Subject No.	Sex	CA	B	IQ**
1	m	7-11	9	57	1	m	6-6	12	93	1	m	7-6	13	111
2	m	6-11	8	75	2	f	6-10	11	89	2	m	6-9	11	96
3	m	7-3	9	60	3	m	6-10	8	85	3	m	7-5	14	91
4	m	6-4	9	81	4	m	8-0	13	100	4	m	6-11	13	101
5	m	7-2	8	63	5	m	6-1	13	98	5	m	7-3	10	107
6	m	7-8	8	63	6	m	6-9	12	95	6	m	7-5	11	96
7	m	6-9	5	69	7	f	6-8	10	85	7	m	7-10	11	107
8	m	7-10	9	76	8	m	6-10	11	98	8	m	7-6	10	109
9	m	6-11	5	64	9	m	6-1	10	82	9	m	7-2	9	121
10	m	7-8	6	73	10	m	6-10	12	98	10	m	6-11	11	97
11	m	6-8	7	62	11	m	7-7	15	101	11	m	6-10	14	122
12	m	8-7	8	71	12	m	7-11	9	79	12	m	6-10	15	106
13	m	6-6	4	65	13	m	7-9	10	92	13	m	6-9	14	98
14	m	6-4	4	59	14	m	6-4	9	104	14	m	7-4	11	105
15	m	7-8	5	60	15	m	8-2	12	100	15	m	7-7	11	108
16	f	6-7	8	67	16	m	7-3	9	86	16	m	8-0	10	105
17	m	7-4	9	83	17	m	7-7	14	123					
18	m	8-5	9	75	18	m	6-3	12	123					
19	m	8-3	5	68	19	f	7-5	11	101					
20	f	8-6	6	53	20	m	8-2	12	97					
					21	m	8-7	15	88					
					22	m	7-7	12	97					
					23	m	7-9	9	89					
					24	m	6-9	8	85					
					25	m	7-4	12	98					
					26	f	6-3	10	95					
Mean		7-3	7.05	67.20			7-2	11.19	95.42			7-3	11.75	105
Range		6-0 to 8-7	4-9	53-83			6-1 to 8-7	8-15	79-123			6-9 to 8-0	9-15	91-122
SD		0-9	1.85	8.10			0-9	1.94	10.49			0-5	1.81	8.57

*IQ = WISC-R Full-Scale IQ
**IQ = Otis-Lennon IQ

m = Male
f = Female
CA = Chronological age in years and months
B = ESPQ Factor B raw score
SD = Standard deviation
MR = Mentally retarded
LD = Learning disabled
NM = Normal

APPENDIX D

COMPUTATION PROCEDURE FOR THE KRUSKAL-WALLIS ONE-WAY ANALYSIS OF VARIANCE RANK SUMS TEST

All of the N Factor B raw scores of the mentally retarded, learning disabled, and normal subjects were first ranked jointly from least to greatest. Average ranks were used when there were ties. The ranks and sums of ranks R_j are presented in Table IV. From the data presented in Table IV, the value of H, corrected for ties, was computed by the formula (Siegel, 1956, p. 188):

$$H = \frac{\frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} - 3(N+1)}{1 - \frac{\Sigma T}{N^3 - N}}$$

where k = number of subject samples

n_j = number of cases in jth subject samples

N = total number of cases in all subject samples = 62

R_u = sum of ranks in jth subject sample

T = $t^3 - t$ (when t is the number of tied observations in a group of scores)

To correct for ties, it was determined how many groups of ties occurred and how many scores were tied in each group. Thus,

t	2	3	2	7	11	3	8	4	4	3
T	6	24	6	140	1320	24	504	60	60	24

There were ten groups of ties. It can be observed that for any value of t, the value T is a constant where $T = t^3 - t$. According to the formula

$$1 - \frac{\Sigma T}{N^3 - N}$$

$$= 1 - \frac{(6+24+6+140+1320+24+504+60+60+24)}{(62)^3 - 62}$$

$$= .9909$$

$$H = \frac{\frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} - 3(N+1)}{.9909}$$

$$H = \frac{\frac{12}{62(62+1)} \frac{(241.5)^2}{20} + \frac{(1021.5)^2}{26} + \frac{(689.5)^2}{16}}{.9909}$$

$$H = \frac{\frac{12}{3906} (2916.1125 + 40133.161 + 29713.14)}{.9909}$$

$$H = \frac{223.54}{.9909}$$

$$H = 225.59$$

TABLE IV

RANKS, SUMS OF RANKS, AND AVERAGE RANKS OF FACTOR B RAW SCORES FOR THE MENTALLY RETARDED, LEARNING DISABLED, AND NORMAL SUBJECTS

MR		LD		NM	
B	r_1	B	r_2	B	r_3
9	(22)	15	(61)	15	(61)
9	(22)	15	(61)	14	(57.5)
9	(22)	14	(57.5)	14	(57.5)
9	(22)	13	(53.5)	14	(57.5)
9	(22)	13	(53.5)	13	(53.5)
9	(22)	12	(47.5)	13	(53.5)
8	(13)	12	(47.5)	11	(39)
8	(13)	12	(47.5)	11	(39)
8	(13)	12	(47.5)	11	(39)
8	(13)	12	(47.5)	11	(39)
8	(13)	12	(47.5)	11	(39)
7	(9)	12	(47.5)	11	(39)
6	(7.5)	12	(47.5)	10	(31)
6	(7.5)	11	(39)	10	(31)
5	(4.5)	11	(39)	10	(31)
5	(4.5)	11	(39)	9	(22)
5	(4.5)	10	(31)		
5	(4.5)	10	(31)	R_3	= 689.5
4	(1.5)	10	(31)	$\overline{R}_3$	= 43.094
4	(1.5)	10	(31)		
		9	(22)		
R_1	= 241.5	9	(22)		
$\overline{R}_1$	= 12.1	9	(22)		
		9	(22)		
		8	(13)		
		8	(13)		
		R_2	= 1021.5		
		$\overline{R}_2$	= 39.288		

MR = Mentally Retarded
LD = Learning Disabled
NM = Normal
B = Factor B raw scores of ESPQ
r = Rank
R = Sum of ranks
$\overline{R}$ = Average rank

APPENDIX E

COMPUTATION PROCEDURE FOR THE DISTRIBUTION-FREE MULTIPLE COMPARISONS TEST

As in the case of the Kruskel-Wallis procedure, all of the N Factor B raw scores of the mentally retarded, learning disabled, and normal subjects are first ranked, jointly, from least to greatest. Average ranks are used in case of ties. Next, the sum of the ranks for each subject sample are divided by the number of scores, n, in each subject sample in order to obtain the average rank, $\bar{R}$, for each sample. The ranks and average ranks for each of the subject samples are presented in Table IV (See Appendix D). The mean rank, $\bar{R}_1$, of the Factor B raw scores for the mentally retarded subjects was 12.1. The mean rank, $\bar{R}_2$, of the Factor B raw scores for the learning disabled subjects was 39.288. The mean rank, $\bar{R}_3$, of the Factor B raw scores for the normal subjects was 43.094.

The next step was to calculate the absolute differences, $\left|\bar{R}_u - \bar{R}_v\right|$, of the average ranks for each of the (u,v) subject sample pairs and the value $z_{(\alpha/[k(k-1)])}\left[\frac{N(N+1)}{12}\right]^{\frac{1}{2}}\left(\frac{1}{n_u} + \frac{1}{n_v}\right)^{\frac{1}{2}}$. The basic calculations are as follows:

Subjects	$\left\|\bar{R}_u - \bar{R}_v\right\|$	$z_{(\alpha/[k(k-1)])}\left[\frac{N(N+1)}{12}\right]^{\frac{1}{2}}\left(\frac{1}{n_u} + \frac{1}{n_v}\right)^{\frac{1}{2}}$
MR-LD	$\|12.100 - 39.288\| = 27.188$	$(3.49)\left[62(63)/12\right]^{\frac{1}{2}}(1/20 + 1/26)^{\frac{1}{2}} = 18.728$
MR-NM	$\|12.100 - 43.094\| = 30.994$	$(3.49)\left[62(63)/12\right]^{\frac{1}{2}}(1/20 + 1/16)^{\frac{1}{2}} = 21.121$
LD-NM	$\|39.288 - 43.094\| = 3.806$	$(3.49)\left[62(63)/12\right]^{\frac{1}{2}}(1/26 + 1/16)^{\frac{1}{2}} = 20.008$

where MR = Mentally Retarded
LD = Learning Disabled
NM = Normal

$\bar{R}_v, \bar{R}_v$ = Mean ranks corresponding to each of the (u,v) pairs of subjects

α = Probability error rate (.001)

k = Number of subject samples (3)

N = Total number of observations (62)

n_u, n_v = Number of observations in each subject sample (MR = 20, LD = 26, NM = 16)