AN EVALUATION OF SEVERAL METHODS OF PREDICTING FULL SCALE IQ FROM THE ITPA

A Thesis

Presented to

the Faculty of the Department of Psychology

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

By

Jean Milroy Humphrey

August, 1971

AN EVALUATION OF SEVERAL METHODS OF PREDICTING FULL SCALE IQ FROM THE ITPA

An Abstract of a Thesis

Presented to

the Faculty of the Department of Psychology

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

By

Jean Milroy Humphrey

August, 1971

ABSTRACT

The purpose of the study was to investigate the usefulness of the ITPA as a predictor of IQ. Three methods of estimating from the ITPA IQ were considered. One of these, the estimate of S-B IQ, is provided by the manual but lacked cross-validation. The other two estimates involve simpler methods which might be used more easily by an ITPA examiner.

The ITPA and the Wechsler Intelligence Scale for Children were administered to 138 children. Two subgroups were also considered, the normative subgroup (IQ=85-115) and the low IQ subgroup (IQ=50-84). For each child, three estimates of IQ were obtained from the ITPA, the estimate of S-B IQ, the SS-derived IQ, and the PLQ. These three estimates were compared to the Wechsler FSIQ. Only the estimate of S-B IQ was not significantly different from FSIQ regardless of IQ range. Correlations between the FSIQ and each of the IQ estimates were substantial and comparable to the reliability reported for the ITPA (.87).

After discussing the difference in light of statistical significance, the empirical significance of the results was considered. It was concluded that for practical purposes any one of the three IQ estimates will serve equally well.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION AND REVIEW OF THE LITERATURE

One of the dilemmas which pervades educational settings is a shortage of personnel for administration of psychological tests and planning educational programs for children with learning difficulties. It is generally admitted that reevaluation of the child's progress occurs too seldom, if at all. The Illinois Test of Psycholinguistic Abilities (ITPA) has been useful in diagnosis of learning difficulties. The purpose of this study is to explore the feasibility of using this test to provide, in addition, an indication of the child's intellectual ability. The ITPA manual provides one method of estimating IQ which as yet lacks cross-validation. Other estimates of IQ derived from the ITPA involving simpler procedures are possible and should be considered as alternative methods to the one proposed by the manual.

The use of the ITPA to estimate IQ would in many cases eliminate the need for an intelligence test and thereby free the psychologist or educational diagnostician to work with other children or to engage in more frequent reevaluation of those previously tested.

Review of the Literature

The ITPA is currently in wide use in educational settings. The Revised Edition (Kirk and McCarthy, 1968) consists of ten primary subtests from which the composite scores are obtained, and two supplementary subtests. The test, constructed after Osgood's model of the communication process (Osgood, 1957), consists of three dimensions: psycholinguistic processes (receptive, expressive and organizing), levels of organization (automatic and representational), and channels of communication (auditory-vocal, auditory-motor, visual-vocal, and visual-motor). Each of the twelve subtests, listed below, may be defined in terms of the three dimensions:

Auditory Reception	Grammatic Closure
Visual Reception	Visual Closure
Auditory Association	Auditory Sequential Memory
Visual Association	Visual Sequential Memory
\erbal Expression	Auditory Closure (Supplementary)
Manual Expression	Sound Blending (Supplementary)

Subtest scores are combined in three major ways to evaluate the child's psycholinguistic abilities (Paraskevopoulos and Kirk, 1969). The composite psycholinguistic age (PLA) provides an estimate of the level of psycholinguistic development expressed in age scores and is based upon the total raw score of the ten subtests.

The psycholinguistic quotient (PLQ) is described as a global index of the child's rate of psycholinguistic development. The PLQ, a ratio of the composite PLA to the chronological age (CA), is obtained by the following computation:

$$PLQ = \frac{PLA}{CA} \times 100$$

The composite scaled score (SS) is obtained by determining the mean of the scaled scores of the ten primary subtests. The composite SS, an indication of the child's performance relative to that of others of the same age, is distributed with a mean of thirty-six and a standard deviation of six. It is stated in the technical manual (Paraskevopoulos and Kirk, 1969) that these values were chosen in order to avoid comparison with the conventional IQ distributions based on a mean of 100 and a standard deviation of fifteen or sixteen. The rationale behind precluding comparison is that the intelligence tests are standardized on a sample including the full range of intellectual ability, while the ITPA normative group is restricted in range of intelligence. In contrast to raw and PLA scores, the SS takes into account the variability of scores about the mean.

In addition to these three measures of psycholinguistic development, the PLA, PLQ, and composite SS, a method of estimating IQ from the total raw score is suggested by Paraskevopoulos and Kirk (1969). To provide an estimate of IQ, a comparison was made between Stanford-Binet (short form) mental ages and the total raw scores of the normative group. Based on this comparison a table of raw scores and corresponding MA estimates was developed. With this estimated MA and the child's CA, the tables provided in the Stanford-Binet manual (Terman and Merrill, 1960) may be entered to obtain an estimate of Stanford-Binet (S-B) IQ. However, many users of the ITPA are not sophisticated in the administration of the Stanford-Binet and may not have the test manual easily available for obtaining the estimate of S-B IQ.

The ITPA has been standardized on a narrowly defined population, which includes only children having average intellectual ability as measured by the Stanford-Binet (short form). Only children having IQs within a range created by the mean plus and minus one standard deviation (Stanford-Binet IQ of 84-116) were included. Such a restricted normative sample is considered by Paraskevopoulos and Kirk (1969) to be the most relevant population in terms of the specific purpose of the test, namely diagnosis of learning disabilities.

Children with learning disabilities are described as those who are unable to learn within the typical classroom situation because of a specific retardation or disorder in one of the processes of speech, language, reading, arithmetic, perception, writing, or behavior. Kirk and Bateman (1962) point out that learning disabilities are those educational problems listed above which are not the result of mental retardation, sensory deprivation, or cultural or instructional factors.

Although the authors of the test state clearly that the test is intended for the identification of learning deficits in otherwise "average" children, the ITPA has been called upon for use with children at all levels of intellectual ability. Jorstad (1971) reports the employment of ITPA results in developing remedial programs for disadvantaged Mexican-American children, the majority of whom were eligible for educable mentally retarded classes. Bateman (1965) supplies two case histories illustrating the usefulness of the ITPA in differential diagnosis and program planning for mentally retarded children. Kirk and Bateman (1962) present a child with an IQ of 83 (Stanford-Binet) in an example of the application of the ITPA for diagnosis of deficits and subsequent remedy of those deficits. Kirk, in a personal communication, also discusses the possibility of a child's being classified as mentally retarded by ordinary mental tests when instead he has the characteristics of learning disability. Such a statement implies that the ITPA may yield useful information about children who on the basis of IQ scores are not represented by the normative sample.

The test has been utilized with other children who likewise are not represented in the normative sample, in particular culturally disadvantaged children (Sigel and Perry, 1968; Howard *et al*., 1970;

Rice and Doughtie, 1970; Jorstad, 1971).

Establishing the validity of the ITPA as a measure of intelligence has been supported by several studies showing the relationship between the ITPA and generally accepted measures of intelligence. The relationship was first suggested by Kirk, McCarthy, and Kirk (1968) in the introduction to the Examiner's Manual: "It is a diagnostic test of specific cognitive abilities, as well as a molar test of intelligence." In the technical manual accompanying the ITPA (Paraskevopoulos and Kirk, 1969) the relationship between the Stanford-Binet (short form) and the ITPA is computed for the 962 children of the normative group. Without correction for restriction of range, the obtained correlations between the Stanford-Binet IQ and composite SS vary from .38 to .64 across the various age groups. The correlations between PLQ and Binet IQ range from .41 to .67 for the various age groups.

McCarthy and Olson (1964) found that mental age (from Stanford-Binet, Form L) and ITPA scores (1961 Experimental Edition) were substantially and positively related. Washington and Teska (1970) using the Experimental Edition of the ITPA in a study of ninety-six disadvantaged children found a correlation of .85 between Binet MA and ITPA PLA.

Adding to the validity of the ITPA as a predictor of IQ are several studies dealing with prediction of achievement from ITPA scores. Washington and Teska (1970) administered the ITPA (Experimental Edition), the Stanford-Binet, Form L, and the California Achievement Tests to ninety-six disadvantaged children in kindergarten through second grades. A correlation coefficient of .81 was reported between the Binet and the total achievement battery. An identical correlation coefficient was obtained between the ITPA and the achievement measure.

Hirshoren (1967) administered the Stanford-Binet and ITPA to sixty-six children (forty Caucasian and twenty-six Negro children) at the beginning of kindergarten. Two years later, at the beginning of the second grade, the children were retested with the ITPA and the California Achievement Tests. The results indicate that the total language score of the ITPA has significant utility as a predictor of all areas of school achievement for the entire group and for both the Caucasian and Negro subgroups. While the Binet shows significant prognostic utility for all achievement areas of the entire group and the Caucasian subgroup, it is useful only for prediction of arithmetic achievement for the Negro subgroup. Hirshoren also reports that including the IQ in a stepwise multiple regression analysis does not increase the prognostic utility beyond that found with the ITPA alone.

Mueller (1969) reports similar testing of a group of 89 children with the Stanford-Binet IQs ranging from fifty to eighty. Mueller administered the ITPA, the Stanford-Binet, and other ability tests and then two years later tested the children with the Wide Range Achievement Tests and the New York Achievement Tests. The correlation, without correction for restricted range, between the Stanford-Binet and total achievement is reported to be .52 and between the ITPA and total achievement to be .53.

In summary, the literature reviewed suggests that the ITPA is now being used for children within a much wider intellectual range than that for which the test was designed. The literature has also shown evidence of both construct and predictive validity which supports the use of the ITPA in estimating intelligence. Cross-validation of the method devised by Paraskevopoulos and Kirk for estimating S-B IQ is indicated for both restricted range IQ and full range samples. Other scores provided by the ITPA were discussed. These include the PLQ and the composite SS from which two additional estimates of IQ may be derived.

Hypotheses

The specific hypotheses to be tested are, that for the entire group which is unrestricted in range of IQ and for both subgroups:

1. There is no significant difference among the four IQs, i.e., Wechsler FSIQ with the three ITPA derived IQs.
2. The correlation among the four IQ distributions will be substantial and empirically significant--of the magnitude of test-retest reliability coefficients.

CHAPTER II

PROCEDURE

Subjects

The sample consists of 138 children, ages four through ten, with a mean age of seven years, six months. The children were selected from two subject pools without regard for age, sex, or ethnic membership, and include those experiencing difficulty in school as well as those progressing normally. Two subgroups are considered within, and in addition to, the entire group. The normative subgroup consists of 74 subjects (IQ 85-115) while the low IQ subgroup (IQ 50-84) includes 46 children with IQs below the normative range.

Tests

The ITPA and the Wechsler Intelligence Scale for Children (WISC) were administered to all children except 26 who received the Wechsler Preschool and Primary Scale of Intelligence (WPPSI). The Wechsler scales were chosen to provide the estimate of intellectual ability because of the current preference for this test over the Stanford-Binet.

Examiners

Testing was completed by a group of first and second year graduate students who were qualified to administer the instruments used. A second group was composed of qualified examiners who were involved in the evaluation of children in several local school systems.

Method

After all tests had been administered and scored, three estimates of IQ based on the results of the ITPA were determined. The estimate of S-B IQ is that suggested in the technical manual and described previously. The second estimate of IQ is simply the PLQ described earlier.

The third estimate was obtained by making a linear transformation of the composite scaled scores to a distribution having a mean of 100 and a standard deviation of 15. Table 1 shows each composite SS with its transformed score equivalent. This estimate is referred to as the SS-derived IQ.

The four IQs, the Wechsler FSIQ and the three ITPA estimates, were compared. These comparisons were made for the entire group and for the two subgroups. When computing correlation coefficients for the normative subgroup, correction was made for restricted intellectual range with the following formula (Gulliksen, 1950):

$$R_{xy} = \sqrt{1 - (1 - r_{xy}^2)\ \frac{s_y^2}{S_y^2}}$$

However, such a correction was not deemed necessary for the low IQ group since the variance of FSIQ for the subgroup approximated that of the standardization group of the WISC.

TABLE 1

TRANSFORMATION OF COMPOSITE SS TO ESTIMATE OF IQ

Composite SS	SS-derived IQ	Composite SS	SS-derived IQ
18	55	37	102
19	58	38	105
20	60	39	108
21	62	40	110
22	65	41	112
23	68	42	115
24	70	43	118
25	72	44	120
26	75	45	122
27	78	46	125
28	80	47	128
29	82	48	130
30	85	49	132
31	88	50	135
32	90	51	138
33	92	52	140
34	95	53	142
35	98	54	145
36	100		

CHAPTER III

RESULTS

The means and standard deviations for the total sample and the two subgroups are presented in Table 2. A Treatments X Subjects design was used for each of the analyses of the ITPA derived IQs. A correlation matrix was obtained for each of the groups analyzed.

Results for Entire Sample

Table 3 presents the results of these analyses of variance. The test among the four IQs is significant at $\alpha = .05$ indicating that not all four estimates of IQ can be considered valid substitutes for the FSIQ. Further analyses indicate that both the estimate of S-B IQ and the SS-derived IQ, when compared to the FSIQ, do not yield significantly different results. However, the analysis of FSIQ and PLQ does result in a significant F-ratio with $\alpha = .05$.

The matrix in Table 4 shows Pearson product-moment correlations .88, .88, and .87 between the FSIQ and estimate of S-B IQ, SS-derived IQ, and PLQ respectively. Correlations among the three ITPA estimates are .97 or higher.

Normative Subgroup (IQ 85-115)

Results of the analyses of variance conducted for this subgroup are shown in Table 3. The analysis performed among the four IQs yields a significant difference between the FSIQ and the SS-derived IQ estimate and between the FSIQ and the PLQ at $\alpha = .05$. The analysis of variance determined for FSIQ and estimate of S-B IQ was not significant

TABLE 2

MEANS AND STANDARD DEVIATIONS
FOR TOTAL SAMPLE AND TWO SUBGROUPS

	Total Sample		Normative Subgroup		Low IQ Subgroup	
	Mean	SD	Mean	SD	Mean	SD
CA	89.73	18.13	92.42	16.88	87.70	22.35
FSIQ	94.69	18.56	100.09	8.34	71.80	13.22
S-B IQ	94.00	20.04	98.89	12.10	72.41	15.77
SS-derived IQ	93.58	15.71	98.26	8.82	75.63	15.76
PLQ	92.54	18.28	96.88	10.94	72.85	15.10

at this level. After correction for restricted range of intelligence, the intercorrelations are .91, .91, and .89 between FSIQ and estimate of S-B IQ, SS-derived IQ and PLQ respectively. The intercorrelations of the three ITPA derived IQ estimates are .94 or higher.

Low IQ Subgroup

The results of the analysis of variance for this subgroup are shown in Table 3. The analysis among the four IQs yielded a significant difference at the .05 level. The analysis between FSIQ and SS-derived IQ was also significant. The analysis of variance for the estimate of S-B IQ and FSIQ and also that for the PLQ and FSIQ yielded no significant difference.

The intercorrelations shown in Table 4 are .81, .84, and .33 for the FSIQ and estimate of S-B IQ, SS-derived IQ, and PLQ, respectively. All three estimates of IQ based upon ITPA scores correlate with each other at .97 or higher.

TABLE 3

F-RATIOS FOR TOTAL SAMPLE AND TWO SUBGROUPS

	Total Sample			Normative Subgroup			Low IQ Subgroup		
	F-Value	df	$F_{.05}$	F-Value	df	$F_{.05}$	F-Value	df	$F_{.05}$
Over all IQ distributions	4.19*	3,120	2.68	6.59*	3,60	2.76	6.16*	3,40	2.84
FSIQ, S-B IQ	.70	1,120	3.92	1.36	1,60	4.00	.20	1,40	4.08
FSIQ, SS-derived IQ	2.14	1,120	3.92	5.12*	1,60	4.00	8.98*	1,40	4.08
FSIQ, PLQ	7.43*	1,120	3.92	10.77*	1,60	4.00	.71	1,40	4.08

TABLE 4

CORRELATION MATRICES

A. Total Sample

	FSIQ	S-B IQ	SS-derived IQ
S-B IQ	.88		
SS-Derived IQ	.88	.97	
PLQ	.87	.99	.97

B. Normative Subgroup*

	FSIQ	S-B IQ	SS-derived IQ	PLQ
FSIQ		.91	.91	.89
S-B IQ	.67			
SS-derived IQ	.67	.95		
PLQ	.65	.99	.94	

*The intercorrelations in the lower half are for restricted intelligence range; those in the upper half are corrected for restricted range of intelligence.

C. Low IQ Subgroup

	FSIQ	S-B IQ	SS-derived IQ
S-B IQ	.81		
SS-derived IQ	.84	.97	
PLQ	.83	.98	.99

CHAPTER IV

DISCUSSION

Looking at the correlation matrix for the total group and for each of the subgroups, all estimates of IQ correlated to about the same degree with the FSIQ. Very high correlations were obtained among the three IQs estimated from the ITPA. These estimates are .94 or higher.

However, the analysis of variance among and between the various IQ estimates is a more critical test.

1. For the full range IQ sample the PLQ as an estimate of IQ was shown to differ from the FSIQ, while the estimate of S-B IQ and SS-derived IQ do not differ significantly from the FSIQ.
2. For the subgroup with the restricted range of IQ comparable to the ITPA normative group, only the estimate of S-B IQ was shown not to differ significantly from the FSIQ.
3. For the low IQ group, neither the PLQ nor the estimate of S-B IQ were significantly different from the FSIQ, but this was not the case for the SS-derived IQ.

It must be noted that the estimate of S-B IQ is the only IQ estimate derived from the ITPA which did not differ significantly from the FSIQ regardless of the range of IQ of the sample studied.

The significant difference found between the SS-derived IQ and Wechsler IQ may be attributed to the error imposed by comparing scores obtained from two different distributions; one covering the full range of IQ (Wechsler) and the other with a limited range of IQ (ITPA).

In a sample normally distributed over the dimension of IQ, about 68 percent of the sample falls within ± 1SD from the mean. This 68 percent represents the normative group for the ITPA. According to the technical manual, ninety-eight percent of these average IQ children (± 1SD Binet IQ) scored within ± 2SD on the ITPA. Thirty-two percent of the subjects in the IQ distribution, i.e., those at the upper and lower extremes of the IQ distribution, are poorly represented in the ITPA distribution. In addition it should be noted that the subjects forming the range of ± 1SD for the IQ distribution form a range of ± 2SD in the ITPA distribution because of restricted IQ range of the ITPA sample.

Therefore, it becomes evident that an estimate which will accurately predict IQ must be one which is not based on an indication of placement within the ITPA distribution, i.e., one which is not based on a standard score.

The estimate of S-B IQ is based on the total raw score obtained on the ITPA. The MA corresponding to this raw score had been determined by comparison of raw scores and MA's which ranged from two years to ten years, ten months. Therefore, this estimate is not subject to the effects of restricted range. Thus, it is not surprising that the estimate of S-B IQ was the only estimate which did not differ from the FSIQ regardless of the range of IQ of the sample studied. These results also provide cross-validation of this method of estimating IQ.

Paraskevopoulos and Kirk (1969) in devising this method of estimating S-B IQ used the restricted normative sample. The results of this study not only provides cross-validation for this method, but also confirm its usefulness for the full range of intellectual ability.

The results of this study should be considered in light of the particular statistics applied to the data. A Treatments X Subjects design is such that with a high correlation among treatments (IQ) it becomes more difficult to obtain significant differences among the means because the design does control for subject variance which is a function of the correlation among the treatments. On the one hand, a high correlation was desired to demonstrate a close relationship among treatments, while knowing that the higher the correlation, the less likely a difference will be statistically significant.

It seems that an empirical decision may also be justified. The correlation coefficients are comparable to the ITPA test-retest reliability of .87. For the entire group, the difference between means for any of the estimates is within two IQ points. For practical purposes it seems that any one of the three methods will estimate equally well. For those who do not have a Stanford-Binet manual easily accessible, the SS-derived IQ and the PLQ are alternative estimates which are quickly and easily derived.

BIBLIOGRAPHY

BIBLIOGRAPHY

Bateman, Barbara. The role of the ITPA in differential diagnosis and program planning for mentally retarded. American Journal of Orthopsychiatry, 1965, 35, 465-472.

Gallagher, James F., and Moss, James W. New concepts of intelligence and their effect on exceptional children. Exceptional Children, 1963, 30, 1-5.

Gulliksen, H. Theory of mental tests. New York: Wiley, 1950.

Hirshoren, Alfred. The prognostic and diagnostic utility and the stability of the Illinois Test of Psycholinguistic Abilities. Unpublished doctoral dissertation, University of Illinois, 1967.

Howard, M. J., Jr., Hoops, H.R., and McKinnon, J. Language abilities of children with differing socioeconomic backgrounds. Journal of Learning Disabilities, 1970, 3, 328-335.

Jorstad, Dorothy. Psycholinguistic learning disabilities in 20 Mexican-American students. Journal of Learning Disabilities, 1971, 4, 143-149.

Kirk, Samuel. The Illinois Test of Psycholinguistic Abilities: Its origin and implications. An unpublished paper relayed by personal communication.

Kirk, Samuel A., and Bateman, Barbara. Diagnosis and remediation of learning disabilities. Exceptional Children, 1962, 29, 73-78.

Kirk, S. A., and McCarthy, J. J. The Illinois Test of Psycholinguistic Abilities--an approach to differential diagnosis. American Journal of Mental Deficiency, 1961, 66, 399-412.

McCarthy, J. J., and Olson, J. L. Validity of studies on the Illinois Test of Psycholinguistic Abilities. Madison, Wisconsin: Photo Press, Inc., Xerlith Service, 1964.

Mueller, M. W. Comparison of psycholinguistic patterns of gifted and retarded children. Selected Convention Papers, 42nd Annual CEC Convention. Washington, D. C.: Council for Exceptional Children, NEA, 1964, 143-148.

Mueller, Max. Prediction of achievement of educable mentally retarded children. American Journal of Mental Deficiency, 1969, 73, 590-596.

Osgood, C. E. Motivational dynamics of language behavior. In M. R. Jones (Ed.), Nebraska symposium on motivation. Lincoln, Neb.: University of Nebraska Press, 1957, 348-424.

Paraskevopoulos, John N., and Kirk, Samuel A. The development and psychometric characteristics of the Revised Illinois Test of Psycholinguistic Abilities. Urbana: University of Illinois Press, 1969.

Rice, James A., and Doughtie, Eugene B. IQ and the ITPA: Classification versus diagnosis. Journal of Learning Disabilities, 1970, 3, 73-78.

Sigel, Irving E., and Perry, Cereta. Psycholinguistic diversity among "culturally deprived" children. American Journal of Orthopsychiatry, 1968, 38, 122-126.

Terman, L. M., and Merrill, Maud A. Stanford-Binet Intelligence Scale: manual for the third revision, form L-M. Boston: Houghton-Mifflin, 1960.

Washington, Ernest D., and Teska, James A. Correlations between the Wide Range Achievement Test, the California Achievement Tests, the Stanford-Binet, and the Illionois Test of Psycholinguistic Abilities. Psychological Reports, 1970, 26, 291-294.