DOCUMENT RESUME

ED 166 215 TM 007 858

AUTHOR Reynolds, Cecil R.; Hartlage, Lawrence C.
TITLE Comparison of WISC and WISC-R Racial Regression Lines.
PUB DATE Mar 78
NOTE 16p.; Paper presented at the Annual Meeting of the Southeastern Psychological Association (24th, Atlanta, Georgia, March 15-18, 1978)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Academic Achievement; Achievement Tests; *Black Students; *Caucasian Students; Comparative Testing; Elementary Education; Intelligence Differences; Intelligence Quotient; *Intelligence Tests; *Predictive Ability (Testing); Predictive Validity; *Racial Differences; Test Bias
IDENTIFIERS Wechsler Intelligence Scale for Children; Wechsler Intelligence Scale for Children (Revised); Wide Range Achievement Test

ABSTRACT

Regression lines for the prediction of Wide Range Achievement Test (WRAT) scores by Wechsler Intelligence Scale for Children (WISC) and WISC Revised (WISC-R) intelligence quotients were compared across race, by the Potoff technique. Subjects were 36 black and 30 white children who were referred for psychological evaluation because of learning or behavior problems. Regression lines for blacks and white did not differ significantly for the prediction of WRAT scores by either the WISC or WISC-R. Results lend general support for the use of a common regression line in the prediction of achievement scores for the two races. (Author/GDC)

Comparison of WISC and WISC-R Racial Regression Lines

Cecil R. Reynolds
Department of Educational
Psychology and Measurement
130 Bancroft Hall
University of Nebraska-Lincoln
Lincoln, Nebraska 68588

Lawrence C. Hartlage
Head, Neuropsychology Section
Dept. of Neurology
Medical College of Georgia
Augusta, Georgia 30901

A paper presented to the 24th annual meeting of the Southeastern Psychological Association, Atlanta: March 15 - 18, 1978.

Abstract

Regression lines for the prediction of WRAT scores by WISC and WISC-R FSIQ's were compared across race, by the Potthoff technique, for 36 black and 30 white referred children. Regression lines for blacks and whites did not differ significantly for the prediction of WRAT scores by either the WISC or WISC-R. Results lend general support for the use of a common regression line in the prediction of achievement scores for the two races.

The use of psychological tests normed primarily with white children for psychological diagnosis and educational decision-making concerning minority children, has come under increasing scrutiny in recent years. Although much discussion of the issue has appeared in both the scientific and public literature, few data of relevance to the issue (with school age children) have been presented. The use of such tests are of special concern to the school psychologist, particularly in view of the Larry P. case and PL 94-142. Harrington (1976, 1975) has gone so far as to state that it is not possible for tests developed and normed on a white majority to be other than biased against minorities and to show less predictive validity when used with minorities.

In response to pressure from the Black Psychological Association (which was actually requesting a moratorium on the use of psychological and educational tests with disadvantaged students), the APA Board of Directors requested, in 1968, the Board of Scientific Affairs to appoint a group to study the use of such tests with disadvantaged students. In reporting on this issue, the

committee (Cleary, Humphreys, Kendrick, and Wesman, 1975) offered a definition of test bias. While including content and construct validity as important variables in the issue of test bias, the focus was clearly on predictive validity:

> A test is considered fair for a particular use if the inference drawn from the test score is made with the smallest feasible random error and if there is no constant error in the inference as a function of membership in a particular group. [Cleary et al., 1975, p. 25]

The definition of bias offered by the APA committee is a restatement of previous definitions by Cardall and Coffman (1964), Cleary (1968), and Potthoff (1966), and has been widely accepted (though certainly not without criticism, e.g., Bernal, 1975; Linn & Werts, 1971; Thorndike, 1971). Oakland and Matuszek (1977) examined class placement procedures under several proposed models of bias and demonstrated that the Cleary model results in the smallest number of children being misplaced, although under certain legislative conditions, Oakland and Matuszek favored the Thorndike (1971) selection model. A statistical technique provided by Potthoff (1966) has also received widespread acceptance in the examination of regression lines to test bias under the Cleary et al. definition.

While considerable data are available on the validity of the SAT for blacks and whites, few studies have appeared dealing with differential validity of IQ tests. Mitchell (1967) studied the validity of two broad based readiness tests to predict first grade achievement for blacks and whites finding similar validity coefficients for the two races. Mitchell's study was limited to comparing the magnitude of independent-dependent variable correlation and did not look for identity of regression lines. Hartlage, Lucas, and Godwin (1976) compared the predictive validity of the WISC and Raven with a group of low SES, disadvantaged children. When comparing what they considered to be the relatively culture-fair test, the Raven Matrices, with the "culture-loaded" 1949 WISC, Hartlage et al. (1976) found the WISC to have consistently larger correlations with measures of reading, spelling, and arithmetic than the Raven. These authors only compared the strength of the relationship in each case and did not look for identity of regression lines (equivalent beta coefficients and intercept constants).

Jensen (1976) and Jensen and Figueroa (1975) investigated the construct and content validity of several popular psychometric instruments across race, including

the WISC-R. Jensen (1976; and Figueroa, 1975) concluded, finding no differences to support claims of bias, the tests evaluated were equally valid for use with blacks, whites, and chicanos. While Jensen's research clearly supports the equivalency of internal psychometric properties of each instrument investigated across race, Jensen neglected to investigate the predictive validity of the various measures for blacks, whites, and chicanos, causing his conclusions to be at best premature. The major purpose of the present study is to begin providing data which will aid in the empirical evaluation of test bias under the Cleary et al. (1975) definition for one of the mostly widely used assessment instruments, the WISC-R, and its predecessor, the 1949 WISC.

Method

Subjects

Thirty white and 36 black school age children referred for psychological evaluation for a wide range of learning/behavior disorders comprised the sample. Sample characteristics are described in detail in table 1.

Insert table 1 about here

Children were selected for inclusion randomly, by a certified school psychologist, from referrals received within a public school system. The psychologist selecting subjects and conducting the evaluations had no knowledge of the purpose of the study.

Procedure

All children were tested on the WISC and WISC-R in a counterbalanced design; the time between testings ranging from six weeks to several months. The Wide Range Achievement Test (WRAT) Reading and Arithmetic subtests were administered at the time of the second testing. Pearson product-moment correlations between the WISC and WISC-R Full Scale IQ's and WRAT Reading and Arithmetic subtests were computed for the total sample and separately by race. Two methods of examining differential validity (or bias) were employed.

The regression lines for each pair of scores were examined by race through the Potthoff (1966) technique. The Potthoff technique is widely accepted and provides a simultaneous test of regression coefficients and intercept values to determine whether two regression lines are equivalent (in this case the black regression line and the white regression line). The magnitude of the correlations

between each set of independent-dependent variables for the two races was examined by testing the difference between the correlations for the two groups. The method chosen is based on Fisher Z - transformations and the standard error of the correlation. The second set of analyses are related to the comparison of regression coefficients encountered in the Potthoff analysis and their computation is somewhat redundant. The second set of analyses is similar to methods used by Mitchell (1967) and Hartlage et al. (1976) and were computed to determine whether different conclusions would have been obtained had only the strength of the relationship between independent-dependent variable pairs been examined across race. Due to the sensitivity of the Potthoff procedure and the number (10) of comparisons being made, a conservative p of .01 was adopted.

Results and Discussion

Regression lines for blacks and whites did not differ significantly for the prediction of WRAT Reading and Arithmetic scores by either the WISC or WISC-R. Regression lines for predicting WISC-R IQ's from WISC IQ's for blacks and whites also showed a high degree of

similarity. Table 2 presents a summary of the results of the F-tests for homogeneity of regression derived from application of Rotthoff's (1966) formulation. The present results are generally consistent with previous

---

Insert table 2 about here

---

investigations of emplbyment tests (Bartlett & O'Leary, 1969; Einhorn & Bass, 1971) and the prediction of college success with the SAT (Cleary, 1968; Kallingal, 1971; Pfeifer & Sedlacek, 1971).

When only the magnitude of independent-dependent variable relationships were examined, as in some previous studies (Hartlage et al. 1976; Mitchell, 1967), no significant difference occurred between the correlations of the WRAT subtests with the WISC and WISC-R Full Scale IQ's for black and white children. Correlations between the two WRAT subtests and between the WISC and WISC-R were also highly similar. These data are summarized in table 2.

Although not dealing with the social issues and consequences surrounding intelligence testing, the present results do provide support for the use of a common

regression line to predict achievement scores for black and white children with the popular WISC and WISC-R. Since homogeneity of regression is commonly assumed in the actual practice of clinical and educational assessment, data such as that presented herein should become increasingly important. Prior to drawing firm conclusions regarding the differential predictive validity of the WISC or WISC-R across race, much more research is needed utilizing a wide variety of criteria, including other individual achievement tests, group achievement measures, teacher-made tests, and teacher ratings of achievement. Perhaps similar studies utilizing sex as the major variable will also be useful.

References

Bartlett, C.J., & O'Leary, B.S. A differential prediction model to moderate the effect of heterogeneous groups in personnel selection and classification. Personnel Psychology, 1969, 22, 1-18.

Bernal, E.M. A response to "Educational uses of tests with disadvantaged subjects." American Psychologist, 1975, 30, 93-95.

Cardall, C., & Coffman, W. A method for comparing the performance of different groups on the items in a test. Research and Development Reports, 1964, 64-5 no. 9, College Entrance Examinations Board.

Cleary, T.A. Test bias: Prediction of grades of negro and white students in integrated colleges. Journal of Educational Measurement, 1968, 5, 115-124.

Cleary, T.A., Humphreys, L.G., Kendrick, S.A., & Wesman, A. Educational uses of tests with disadvantaged students. American Psychologist, 1975, 30, 15-41.

Einhorn, H.J., & Bass, A.R. Methodological considerations relevant to discrimination in employment testing. Psychological Bulletin, 1971, 75, 261-269.

Harrington, G.M. Minority test bias as a psychometric artifact: The experimental evidence. paper presented at the symposium, Race and sex differences in ability, annual meeting of the American Psychological Association, Washington, September 1976.

Harrington, G.M. Intelligence tests may favour the majority groups in a population. Nature, 1975, 258, 708-709.

Hartlage, L.C., Lucas, T., & Godwin, A. Culturally biased and culture fair tests correlated with school performance in culturally disadvantaged children. Journal of Clinical Psychology, 1976, 32, 235-237.

Jensen, A.R. Test bias and construct validity. Phi Delta Kappan, 1976, 58, 340-346.

Jensen, A.R., & Figueroa, R.A. Forward and backward digit-span interaction with race and IQ. Journal of Educational Psychology, 1975, 67, 882-893.

Kallingal, A. The prediction of grades for black and white students at Michigan State University. Journal of Educational Measurement, 1971, 8, 263-265.

Linn, R.L., & Werts, C.E. Considerations for studies of test bias. Journal of Educational Measurement, 1971, 8, 1-4.

Mitchell, B.C. Predictive validity of the Metropolitan Readiness Tests and the Murphy-Durrell Reading Readiness Analysis for white and for negro pupils. Educational and Psychological Measurement, 1967, 27, 1047-1054.

Oakland, T., & Matuszek, P. Using tests in non-discriminatory assessment. in T. Oakland (Ed.) Psychological and Educational Assessment of Minority Group Children, New York: Brunner/Mazel, Inc., 1977.

Pfeifer, C.M., & Sedlacek, W.E. The validity of academic predictors for black and white students at a predominately white university. Journal of Educational Measurement, 1971, 8, 253-261.

Potthoff, R.F. Statistical aspects of the problem of biases in psychological tests. Institute of Statistics Mimeo Series No. 479, Chapel Hill, N.C.: UNC-Chapel Hill Department of Statistics, 1966.

Thorndike, R.L. Concepts of culture-fairness. Journal of Educational Measurement, 1971, 8, 63-70.

Table 1

Sample Characteristics by Race

	N	F	M	$\bar{X}$ Age	SD Age	$\bar{X}$ WISC-IQ	IQ SD	$\bar{X}$ WISC-R	IQ SD
Blacks	36	19	17	12.6	1.3	84.41	12.11	79.75	10.68
Whites	30	13	17	12.9	1.8	83.63	19.37	79.00	19.87
Total	66	32	34	12.7	1.5	84.06	15.70	79.41	15.42

Table 2

Summary tabulation of F-tests for similarity of black and white regression lines and comparisons of correlational values in the prediction of achievement.

Tests of Regression Lines	F (2,62)	P
WISC FSIQ - WRAT Reading	1.51	N.S.
WISC-R FSIQ - WRAT Reading	2.17	N.S.
WISC FSIQ - WRAT Arithmetic	2.75	N.S.
WISC-R FSIQ - WRAT Arithmetic	4.05	N.S.
WISC FSIQ - WISC-R FSIQ	1.11	N.S.

Tests for Significance of Difference Between Correlations

	r_b	r_w	Z	P
WISC FSIQ - WRAT Reading	.62	.71	0.44	N.S.
WISC-R FSIQ - WRAT Reading	.40	.77	1.71	N.S.
WISC FSIQ - WRAT Arithmetic	.64	.63	0.05	N.S.
WISC-R FSIQ - WRAT Arithmetic	.69	.72	0.15	N.S.
WISC FSIQ - WISC-R FSIQ	.81	.95	1.02	N.S.