The Reliability of the K-ABC for Hispanic and White Children: A Comparison by Year

Arthur E. Hernandez Victor Willson

Texas A & M University

Running Head: K-ABC Reliability

The Reliability of the K-ABC for Hispanic and White Children:

A Comparison by Year

Abstract

Reliability estimates were calculated for two groups (White and Hispanic) at eleven age levels on the Kaufman Assessment Battery for Children Mental Processing Scales using computed variance components. These estimates were compared as ratios of error variances and evaluated using F distributions. Ratios for the two groups at all age levels proved to be nonsignificant, indicative of statistically equivalent reliabilities for the two groups. Coefficients were, with two exceptions, similar in magnitude. While not a proof of "non-bias", this demonstation supports the notion that error in this instance is a function that is constant across groups.

The Kaufman Assessment Battery for Children (K-ABC) is an individually administered assessment battery designed to measure the intelligence and achievement of 2 1/2 to 12 1/2 year old children (Kaufman and Kaufman, 1983). The K-ABC is intended for, among other things, minority group assessment. Among the psychometric characteristics which must be considered in light of the instrument's claim to validity in minority assessment, must be the estimates of reliability for each group, these should be compared to determine if significant differences exist (Reynolds, 1982). This study was undertaken to compare internal consistency reliability estimates for Hispanics and Whites to ascertain the presence or absence of significant differences.

Method

Subjects

Data were taken from the standardization sample for the K-ABC (Kaufman and Kaufman, 1983). The test was normed on a sample of 2,000 children stratified within each age group by geographic region, socioeconomic status, race or ethnic group, community size and educational placement of the child. Sampling was based on early reports and projections of the 1980 U.S. Census results for all variables except sex. An equal number of children of each sex were sampled at every age level. The N for the White group was 1,569 and for the Hispanic group was 161 comprising

approximately 72 and 8 percent of the total sample respectively. The N for each age group for the White sample ranged from 187 at age 4 to 80 at the 12 year old group with a median size of 143. The N for each age group for the Hispanic sample franged from 19 at the 9 and 10 year age levels to 6 at the 12 1/2 year age level, with a median size of 18.

Procedure

White and Hispanic members of the standardization group of the K-ABC were grouped into eleven whole year age groups. Reliability was estimated using variance components computed by the program VARCOMP of SAS (1979) and following an application of the generalizibility theory (Cronbach, Gleser, Nanda, and Rajaratnam, 1972) where:

$$r_{xx} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_e^2}$$

The resultant reliability coefficients were compared as ratios of error variances, $(1 - r_{xx})$ with the larger variance always over the smaller. Feldt (1969) showed this ratio is distributed as a central F with N - 1 degrees for freedom in the numerator and N - 1 in the denominator when the no difference hypothesis is true. These F values were then evaluated for statistical significance to determine the existence of potential bias in the K-ABC's measurement of the target trait for both groups.

Results

Not one F was statistically significant. However the F for age group 2 did approach significance. The results from this age group must be interpreted carefully since the Hispanic group is extremely small, (N=4) and resultant r_{xx} for this group is somewhat discrepent with the values obtained at other age levels. A comparison of the mean value of r_{xx} for both groups indicates no difference ($r_{xx\ White}$ = .787, SD =.034, $r_{xx\ Hispanic}$ = .736, SD =.178) Removing the one group which is suspect the mean values are quite close.

Insert Table 1 about here

Discussion/Conclusions

For any assessment instrument to be "non-biased" it must show consistency across groups. If bias is defined as systematic as opposed to random error, then it would appear logical to conclude that the lack of significantly different error variances for different groups would be indicative of the lack of any systematic differences in the performance of the measure. However, this should not be construed as a definitive demonstration of non-bias, the small N is for the Hispanic groups suggest that caution is warranted in the interpretation of the results. Bias for this study is defined as a systematic

difference in the ratio of error variances, and should be interpreted as such. This procedure seems important and valid for ruling out a specific aspect of bias, but does not ensure the non-biased nature of the instrument. As a preliminary indication, then, this demonstration suggests that the test is equally reliable for both white and Hispanic children at all age levels.

TABLE 1

	WHITE	HISPANIC	F-RATIO
	.759	.856	1.127
	.752	.231	3.256
	.766	.835	1.094
	.818	.872	1.067
	.800	.742	1.079
	.772	.825	1.069
	.754	.731	1.031
	.812	.773	1.051
	.762	.825	1.082
	.856	.693	1.235
	.807	.719	1.123
Means	.787	.736	
SD	.034	.178	

Bibliography

Cronbach, L. J., Gleser, G.C., Nanda, H., and Rajaratnam, N. (1972) The dependability of behavioral measurements: Theory of generalizability for scores and profiles. New York: Wiley.

Feldt, L. S. (1969) A test of the hypothesis that Cronbach's alpha or Kuder-Richardson coefficient twenty is the same for two tests. Psychometrika, 34, 363-373.

Kaufman, A.S. & Kaufman, N.L. (1983) The Kaufman Assessment Battery for Children. Circle Pines, MN: American Guidance Service.

Reynolds, C. R. (1982) Methods for detecting construct and predictive bias. In R. A. Berk (Ed.) Handbook of methods for detecting test bias. Baltimore: John Hopkins University Press.

Statistical Analysis System, Version 79.6, Cary, NJ: SAS Institute, Inc. 1979.