DOCUMENT RESUME

ED 219 429 TM 820 444

AUTHOR Wolfle, Lee M.; Robertshaw, Dianne
TITLE Racial Differences in Measurement Error in Educational Achievement Models.
PUB DATE Mar 82
NOTE 20p.; Paper presented at the Annual Meeting of the American Educational Research Association (66th, New York, NY, March 19-23, 1982).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; *Analysis of Covariance; Black Students; *Error of Measurement; *Estimation (Mathematics); Factor Structure; Least Squares Statistics; Longitudinal Studies; *Parent Background; *Racial Factors; Research Problems; Secondary Education; Statistical Bias; Statistical Significance; White Students
IDENTIFIERS *Covariance Structural Analysis; Joreskog (Karl G); Occupational Status

ABSTRACT

Comparisons between whites and blacks in models of educational achievement were found to be suspect when based solely on least-square estimates, since the estimates are biased by measurement error varying by race. In this study, white high school seniors were shown to report their parents' status characteristics more reliably than black high school seniors. Data were drawn from the National Longitudinal Study of the High School Class of 1972 as the seniors moved into early adult years. Variables included father's and mother's educational attainment in the base year and first followup year, and father's occupation measured by Duncan's socioeconomic index adjusted to the 1970 census occupation classification. Simultaneous covariance structure analysis methods developed by Joreskog were used in estimating measurement parameters for both groups. Explicit statistical tests of group differences in measurement error patterns are described. (Author/CM)

ED219429

RACIAL DIFFERENCES IN MEASUREMENT ERROR IN EDUCATIONAL ACHIEVEMENT MODELS*

Lee M. Wolfle

Dianne Robertshaw

Virginia Polytechnic Institute and State University

*This paper was prepared for delivery at the annual meetings of the American Educational Research Association, New York, March 19-23, 1982.

TM 820 444

RACIAL DIFFERENCES IN MEASUREMENT ERROR IN EDUCATIONAL ACHIEVEMENT MODELS

ABSTRACT

Racial differences in the accuracy of reports of parental status characteristics are investigated using Joreskog's (1971a) general framework for simultaneous covariance structure analyses of multiple populations. Results indicate that white's reports of these characteristics are significantly more reliable than are those of blacks.

RACIAL DIFFERENCES IN MEASUREMENT ERROR IN EDUCATIONAL ACHIEVEMENT MODELS

Reliable and theoretically meaningful measurement is a prerequisite for good educational research; yet it has not always assumed the prominence it deserves. The linear statistical model is probably the most commonly used analytic tool in educational research, but using the linear statistical model carries with it a clear, but often overlooked, assumption about measurement error. As Blalock (1964, p. 49) noted, one assumes that "there may be errors of measurement with respect to the dependent variable Y, but that all of the independent variables have been measured without error." Such an assumption is obviously unrealistic for most social data, and has a well-known effect upon the least-squares estimators -- they are biased (Walker and Lev, 1953, p. 306).

Until recently, there was little that an educational researcher could do about least-squares indicators biased by measurement error. There were basically only three alternatives. By far the most common was to naively assume that the variables were measured without error, and wistfully hope the resulting estimates were robust. A second alternative was to correct correlation coefficients for attenuation, and use the corrected estimates as inputs to a regression analysis. This procedure, however, required a priori knowledge of the reliability coefficients for the variables; furthermore, one had to assume the reliabilities were invariant from one population, subpopulation, or sample to the one at hand. These restrictions have severely limited the use of regression analyses based on correlations corrected for attenuation.

Yet a third alternative was to measure implied coefficients between latent variables for which one had multiple manifest indicators. Siegel and Hodge (1968), for example, explicated several such models in their paper directed to sociologists; furthermore, they noted that correlations corrected for attentuation were merely special cases of their multiple indicator models. The problem with this alternative, as noted by both Hauser and Goldberger (1971) and Long (1976), is its casual approach toward statistical estimation and hypothesis testing. The problem results from overidentified models, which yield multiple estimates of the associations among latent variables. In response, some authors have chosen to ignore one or more of the identifying equations (e.g., Blalock, 1970; Land, 1970); others have averaged the estimates from the several equations (e.g., Hauser, 1970). A better alternative would be to obtain estimates of the overidentified parameters by maximum likelihood estimation (MLE). These procedures grew out of the work of Lawley (1943), but the immense computational load required for the iterative estimation of maximum-likelihood estimates prevented their application in practice. Thus, the application of more adequate statistical procedures languished until Joreskog (1966, 1967, 1969) discovered an efficient MLE computational procedure, soon to be followed by a computer program for confirmatory factor analysis (Joreskog, Gruvaeus, and van Thillo, 1970). The resulting variances and covariances of the latent factors could be used to estimate the parameters of a structural model assumed to exist among the factors, and Joreskog and Sorbom (1978) have provided a computer program for both a confirmatory factor analysis measurement model, and an hypothesized linear structural model among the factors.

This program is called LISREL, an acronym for linear structural relationships, and possesses the potential for revolutionizing the way educational researchers test hypothesized relationships among theoretical, unmeasured latent variables (Kerlinger, 1977).

These analytic developments have direct application to substantive problems in educational research. Of particular importance, models of educational achievement often include measures of socioeconomic background. Otherwise, outcomes attributed to treatments within schools, for example, could be due to differences in ocioeconomic background. If the socioeconomic indicators are measured with error, however, their least-squares estimates will be biased. And when models are compared across populations, or across different cohorts of the same population, biases can be compounded by differential kinds and magnitudes of measurement error in the several populations.

Several studies suggest that measurement errors in socioeconomic background variables differ between blacks and whites. For example, Bielby et al. (1977) estimated response error models for measures of socioeconomic attainment in black and nonblack populations. Their results suggested that nonblacks report socioeconomic variables with random errors, but blacks report such variables with greater consistency than warranted in reality. That is, if blacks overstate, say, their father's education they are also likely to overstate their mother's. Wolfle and Lichtman (1981) did not find significant nonrandom measurement errors among either whites or blacks, but they did find that blacks reported status levels of their parents with greater errors than did whites. These differences were not trivial in their impact on structural parameter estimates; Patteson and Wolfle (1981) found that least-square regression estimates could be biased by as much as 100 percent when compared to estimates adjusted for the presence of measurement error.

These comparisons of measurement error between blacks and whites are deficient, however, because the authors estimated models independently for each group. Estimating measurement models independently for each group is deficient for at least two reasons. First, it precludes statistical tests of group differences in measurement model parameters. Second, where differences are not found across groups, better parameter estimates are obtained by equating coefficients across groups, thus increasing the number of observations while using fewer degrees of freedom.

A statistical model appropriate for comparing measurement model parameter estimates across groups was proposed by Joreskog (1971a), but to date has received little use in educational research. Mare and Mason (1980) have applied Joreskog's (1971a) analytic framework to children's reports of parental characteristics involving white children in the sixth, ninth and twelfth grades. Corcoran (1980) has studied sex differences in measurement errors for socioeconomic variables. But no one has used Joreskog's (1971a) framework to investigate measurement error differences between blacks and whites. This paper makes explicit statistical tests of measurement model similarities and differences between blacks and whites for variables typically found in models of educational achievement.

THE DATA

Data for this study were drawn from the National Longitudinal Study of the High School Class of 1972 (see Riccobono, et al., 1981). The NLS was designed to provide data on a large cohort of high school seniors, and to follow these students as they made the move from high school into their early years of adulthood. The variables used in this analysis include respondent's reports of their father's and mother's educational attainment in both the base-year (1972) and first follow-up (1973) surveys, and father's occupation measured at the same times in terms of Duncan's (1961) socioeconomic index as adjusted to the 1970 census occupational classification. The analysis reported here is restricted to 647 black respondents who possessed complete reports for these six variables, and 650 white respondents selected at random from the nearly 8000 whites with complete data.

STATISTICAL MODELS

For each race, the basic measurement model can be described by the following six equations:

$$FAED_1 = \lambda_{11}FAEDTRUE + \delta_1$$

$$FAED_2 = \lambda_{21}FAEDTRUE + \delta_2$$

$$MAED_1 = \lambda_{32}MAEDTRUE + \delta_3$$

$$MAED_2 = \lambda_{42}MAEDTRUE + \delta_4$$

$$FAOC_1 = \lambda_{53}FAOCTRUE + \delta_5$$

$$FAOC_2 = \lambda_{63}FAOCTRUE + \delta_6$$

For each true parental status characteristic this is at least a congeneric measures model (Joreskog, 1971b). True scores were allowed to covary, and were not constrained to be equal for blacks and whites. Covariances among response errors were initially set at zero on the assumption that response errors were random, but were subsequently allowed to covary. The statistical tests available for selecting the best-fitting model consist of (1) estimating a model in which certain parameters are set to be equal, sometimes within a racial group, and sometimes across the two groups; and (2) estimating a less constrained model. The test consists of assessing the statistical significance of the improvement in fit going from the more constrained model to the less constrained model. If the more constrained model fits the data as well as the less constrained model (i.e., within sampling error limits), then one may conclude that the constraints do not seriously erode the fit of the model. Suppose, for example, that one model is specified such that the reliabilities of father's education are constrained to be equal for whites and blacks; another model is specified such that these reliabilities are free to vary between races. If the constrained model provides just as good a fit as the free model, then constraining equal reliabilities does not erode the fit of the model to the data, and one would conclude that the reliabilities of father's education were equal for whites and blacks.

RESULTS

The strategy of analysis we followed was to specify a series of measurement models reflecting alternative assumptions about the patterns of errors with which black and white high school seniors reported the status characteristics of their parents. In matrix algebra notation, the model being tested is:

$$X = \Lambda F + \delta$$

in which X is a 6 x 1 vector of reported parental status characteristics, Λ is a 6 x 3 matrix of coefficients, F is a 3 x 1 vector of true parental status characteristics, and δ is a 6 x 1 vector of disturbances. The covariance matrix for the status reports implied by the model is then:

$$\Sigma = \Lambda \Phi \Lambda + \Theta_{\delta}$$

The simplest possible pattern of errors is to assume that all errors were random. In this case, theta-delta, the variance-covariance matrix of disturbances, was specified to be a diagonal matrix, implying that all error covariances were zero. All free coefficients were allowed to vary both within and across groups. The two variance-covariance matrices of observed reports for blacks and whites contain 21 nonredundant coefficients, and within groups there are 15 coefficients being estimated -- six in lambda, three in phi, and six in theta-delta. The variances in phi were fixed at unity to provide standardized estimates.

Table 1 presents goodness-of-fit statistics for this and several other models. Model A of Table 1 is the model of random errors just described. The likelihood ratio chi-square value of the model of random errors was 74.51 with 12 degrees of freedom, indicating that random reporting errors were very unlikely to have generated the observed covariance matrices.

Since errors were not random, it became necessary to consider several forms of nonrandomness. One form of nonrandomness is to assume in the NLS base-year survey (1972) that reporting errors of father's and mother's education were correlated. This could occur, for example, if a respondent knew one parent's educational attainment but not the other's, and guessed the unknown with reference to the known. Another form of nonrandomness could occur if respondents reported their father's educational and occupational status with more consistency than warranted by the truth. Model B in Table 1 was estimated with the covariances between the disturbances of mother's and father's education, and father's education and occupation, in the base-year allowed to be nonzero and estimable by the model for both whites and blacks. The likelihood ratio chi-square was 11.02 with 8 degrees of freedom. This represents a major improvement in the fit of the model, and indicates that nonrandom reporting errors of parental status variables were characteristic of this sample of high school seniors.

Of course, it was also possible for parental status variables to be reported with nonrandom errors in the first follow-up survey (1973). Model C in Table 1 represents such a model, and is also a plausible alternative to a model of random errors. It is even possible that correlated errors existed in both surveys, and Model D reports these results. The fit of this model to the observed data is very good, and represents a point of departure for considering whether reporting errors of parental status characteristics were equal for blacks and whites.

Model E in Table 1 was estimated with all elements (as specified in Model D) in lambda and theta-delta specified to be equal for both blacks and whites. These constraints are equivalent to specifying that the reliability coefficients for whites and blacks are equal. The likelihood ratio chi-square value for this model was 125.73 with 20 degrees of freedom, indicating that whites and blacks do not report parental status characteristics with equal reliabilities.

To this point we have assessed the adequacy of several measurement models merely by examining their fit to the observed data. But we are also interested in finding, if possible, a parsimonious model. Toward this end, we examined the standard errors of the error covariances in Model D (not shown here), and found that all but two of the error covariances were not significantly different from zero. The significant covariances were between the respondent's errors in reporting father's and mother's education in the first follow-up survey for both blacks and whites. Model F was suggested by these observations. In Model F the covariances between the disturbances of mother's and father's education in the follow-up survey for both blacks and whites were allowed to be free, but all other error covariances were fixed at zero. As one would expect, Model F does not fit as well as Model D, but the difference in the likelihood ratio chi-square values, which itself is distributed as chi-square, is not statistically significant (the probability of a chi-square value of 8.30 with 6 degrees of freedom is .217). This result indicates that Model F is to be preferred over Model D; the two models do not differ significantly in fit, but Model F uses fewer degrees of freedom.

Why the disturbances of mother's and father's education were correlated in the first follow-up survey, but not the base-year survey, is a bit of a mystery. The method of data collection, however, differed between the base-year and follow-up surveys, and may have contributed to the anomaly. In the base-year the respondents were given a questionnaire in school, and then given the option of completing the questionnaire in school or taking it home and answering the questions with the assistance of their parents. In contrast, the follow-up survey was conducted primarily by mail, and may not have afforded the respondents the opportunity to resolve uncertainties about their parents' education. In any event, the reporting errors of mother's and father's education were correlated in the follow-up survey.

We next considered whether the reliabilities of the observed status variables were equal for the two surveys. To effect this analysis, we respecified Model F such that the lambda coefficients and the error variances in theta-delta within racial groups were constrained to be equal for each of the two indicators for each of the latent true statuses. For example, the reliability of father's education observed in the base-year was constrained to be equal to the reliability of father's education observed in the follow-up. No constraints were made across either latent status variables or racial groups. Model G in Table 1, which reports this analysis, provides a more parsimonious fit than Model F. Although the chi-square value for Model G is larger than that of Model F, Model G uses fewer degrees of freedom; the difference in chi-square values is not statistically significant.

The reliability coefficients in Model G (not shown here) suggested the possibility that mother's and father's educations were reported by the NLS respondents with equal reliability. The fit of Model H, however, which specified these equalities separately for both whites and blacks, is not very good. But Model I, which specifies equal reliabilities for mother's and father's education for blacks, but not for whites, provides an excellent fit to the observed data. Model I is a parsimonious model that fits the data well, and became our final measurement error model of parental status characteristics.

Parameter and reliability estimates for Model I are shown in Table 2. First, these estimates indicate that whites report their parents' status characteristics more reliably than do blacks. Blacks are only about 90 percent as reliable respondents of parental status as whites. Second, the estimates show that all respondents reported their parents' education more accurately than they reported their father's occupation. Father's occupation was reported only about 80 percent as reliably as either father's or mother's education. Finally, blacks do not favor either parent in reporting their educational attainment, but whites report their father's education slightly more accurately than their mother's.

The middle panel of Table 2 shows the correlation coefficients among the three parental status true scores. These correlations among the latent or true status characteristics are all larger in value than their counterparts among the observed variables. This is not unexpected, as the observed variables contain considerable portions of random (and some nonrandom) error. Because blacks report these variables less reliably than whites, the correlations among the observed variables for blacks are considerably more attenuated than for whites. This has serious consequences for substantive conclusions to be drawn from least-squares estimates of parameters in models of educational achievement. In particular, when comparisons are made between whites and blacks, the least-squares parameter estimates will vary as a function of differential measurement error even if the true effects are the same for both groups.

The final panel of Table 2 presents the estimated correlations between errors in reports of mother's and father's schooling for whites and blacks. These correlations are substantial, and suggest that the NLS respondents surveyed by mail questionnaire resolved their uncertainty about parental schooling by reconciling their parents' schooling.

CONCLUSION

This paper has illustrated the use of simultaneous factor analytic methods developed by Joreskog (1971a) in estimating measurement models of parental status characteristics as reported by white and black high school seniors. By estimating measurement parameters for both groups simultaneously, we have been able to make explicit statistical tests of group differences in measurement error patterns. These tests indicate that whites report their parents' status characteristics more reliably than blacks. These results suggest that comparisons between whites and blacks of models of educational achievement are suspect if they are based solely on least-squares estimates. Such estimates are biased by the presence of measurement error; because measurement error varies by race, interracial comparisons of least-squares structural parameter estimates are biased.

Table 1. Goodness-of-Fit Statistics for Measurement Models for Black and White 1972 High School Graduates

	Model	X^2	Degrees of Freedom	Prob.
A.	Random measurement errors	74.51	12	.000
B.	Covariances among errors of father's and mother's education, and father's education and occupation in base-year survey	11.02	8	.200
C.	Covariances among errors of father's and mother's education, and father's education and occupation in followup survey	7.25	8	.510
D.	Covariances among errors of father's and mother's education, and father's education and occupation in base-year and followup surveys	1.99	4	.738
E.	Model D with black and white measurement error structures constrained to be equal	125.73	20	.000
F.	Covariances among errors of father's and mother's education in followup survey	10.29	10	.415
G.	Model F with reliabilities of father's education equated, mother's education equated, and father's occupation equated	17.49	22	.736
H.	Model G with reliabilities of father's and mother's education equated to each other	32.46	26	.178
I.	Model G with reliabilities of father's and mother's education equated to each other for blacks only	17.50	24	.826

Table 2. Model I Parameter Estimates

Characteristics	Manifest Measure	True Score Variance	Error Variance	Slopes	Reliability
Whites					
Father's education	Base-year Followup	1.00	.088 .088	.955 .955	.916 .916
Mother's education	Base-year Followup	1.00	.119 .119	.940 .940	.888 .888
Father's occupation	Base-year Followup	1.00	.270 .270	.854 .854	.760 .760
Blacks					
Father's education	Base-year Followup	1.00	.175 .175	.910 .910	.839 .839
Mother's education	Base-year Followup	1.00	.175 .175	.910 .910	.839 .839
Father's occupation	Base-year Followup	1.00	.390 .390	.780 .780	.667 .667

True Score Correlations *

	1.	2.	3.
1. Father's Educ.	1.00	.635	.738
2. Mother's Educ.	.594	1.00	.495
3. Father's Occup.	.730	.468	1.00

* Whites below the diagonal, and blacks above.

Correlations between Errors in Reports of Mother's and Father's Education in First Followup

	Correlations
Whites	.322
Blacks	.400

REFERENCES

Bielby, W. T., Hauser, R. M., and Featherman, D. L. "Response errors of black and nonblack males in models of the intergenerational transmission of socioeconomic status." American Journal of Sociology, 1977, 82, 1242-1288.

Blalock, H. M., Jr. Causal Inferences in Nonexperimental Research. Chapel Hill: University of North Carolina Press, 1964.

Blalock, H. M., Jr. "Estimating measurement error using multiple indicators and several points in time." American Sociological Review, 1970, 35, 101-111.

Corcoran, M. "Sex differences in measurement error in status attainment models." Sociological Methods and Research, 1980, 9, 199-217.

Duncan, O. D. "A socioeconomic index for all occupations." In A. J. Reiss (ed.), Occupations and Social Status. New York: The Free Press of Glencoe, 1961.

Hauser, R. M. "Educational stratification in the United States." Sociological Inquiry, 1970, 40, 102-129.

Hauser, R. M., and Goldberger, A. S. "The treatment of unobservable variables in path analysis." In H. L. Costner (ed.), Sociological Methodology 1971. San Francisco: Jossey-Bass, 1971.

Joreskog, K. G. "UMFLA: A computer program for unrestricted maximum likelihood factor analysis." Research Memorandum 66-20. Princeton, NJ: Educational Testing Service, 1966.

Joreskog, K. G. "Some contributions to maximum likelihood factor analysis." Psychometrica, 1967, 31, 165-178.

Joreskog, K. G. "A general approach to confirmatory maximum likelihood factor analysis." Psychometrica, 1969, 34, 183-202.

Joreskog, K. G. "Simultaneous factor analysis in several populations." Psychometrica, 1971a, 36, 409-426.

Joreskog, K. G. "Statistical analysis of sets of congeneric tests." Psychometrica, 1971b, 36, 109-133.

Joreskog, K. G., Gruvaeus, G. T., and van Thillo, M. ACOVS: A General Computer Program for Analysis of Covariance Structures. Princeton, NJ: Educational Testing Service, 1970.

Joreskog, K. G., and Sorbom, D. LISREL IV: Analysis of Linear Structural Relationships by the Method of Maximum Likelihood. Chicago: National Educational Resources, 1978.

Kerlinger, F. N. "The influence of research on educational practice." Educational Researcher, 1977, 6, 5-12.

Land, K. C. "On the estimation of path coefficients for unmeasured variables from correlations among observed variables." Social Forces, 1970, 48, 506-511.

Lawley, D. N. "The application of the maximum likelihood method to factor analysis." British Journal of Psychology, 1943, 33, 172-175.

Long, J. S. "Estimation and hypothesis testing in linear models containing measurement error: A review of Joreskog's model for the analysis of covariance structures." Sociological Methods and Research, 1976, 5, 157-206.

Mare, R. D., and Mason, W. M. "Children's reports of parental socioeconomic status: A multiple group measurement model." Sociological Methods and Research, 1980, 9, 178-198.

Patteson, B. J., and Wolfle, L. M. "Specification bias in causal models with fallible indicators." Multiple Linear Regression Viewpoints, 1981, 11, 75-89.

Riccobono, J., Henderson, L. B., Burkheimer, G. J., Place, C., and Levinsohn, J. R. National Longitudinal Study: Base Year (1972) through Fourth Follow-Up (1979) Data File Users Manual. Washington, D.C.: National Center for Education Statistics, 1981.

Siegel, P. M., and Hodge, R. W., "A causal approach to the study of measurement error." In H. M. Blalock, Jr., and A. B. Blalock (eds.), Methodology in Social Research. New York: McGraw-Hill, 1968.

Walker, H. M., and Lev, J. Statistical Inference. New York: Holt, Rinehart and Winston, 1953.

Wolfle, L. M., and Lichtman, M. "Educational attainment among whites, blacks, and Mexican-Americans." Paper presented to the American Educational Research Association, Los Angeles, 1981.