DOCUMENT RESUME

ED 247 276 TM 840 450

AUTHOR	Wolfle, Lee M.; Ethington, Corinna A.
TITLE	Within-Variable, Between-Occasion Error Covariances in Models of Educational Achievement.
PUB DATE	Apr 84
NOTE	22p.; Paper presented at the Annual Meeting of the American Educational Research Association (68th, New Orleans, LA, April 23-27, 1984).
PUB TYPE	Speeches/Conference Papers (150) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Error of Measurement; Estimation (Mathematics); *Goodness of Fit; High Schools; *Mathematical Models; Maximum Likelihood Statistics; Multivariate Analysis; Parent Background; Research Problems; *Socioeconomic Background; Statistical Bias; White Students
IDENTIFIERS	High School and Beyond (NCES); *LISREL Computer Program

ABSTRACT

To correct for the effects of measurement error on structural parameter estimates, many researchers are now estimating models of educational achievement with LISREL. In order to estimate such models it is desirable to obtain multiple manifest measures of the latent constructs. Many researchers restrict their models to two manifest measures per latent construct for reasons of economy, but doing so assumes, in the absence of external information, that all of the covariance between the within-variable measures is reliable covariance. Such an assumption may or may not hold in practice. The present study empirically investigated the extent of within-variable, between-occasion error covariances among nine socioeconomic variables typically included in models of educational achievement using data on 1064 white respondents from High School and Beyond. Little evidence was found to support the claim that reliability estimates for social background variables are inflated due to correlated errors of measurement. (Author/BS)

WITHIN-VARIABLE, BETWEEN-OCCASION ERROR COVARIANCES IN MODELS OF EDUCATIONAL ACHIEVEMENT

Lee M. Wolfle

and

Corinna A. Ethington

Virginia Polytechnic Institute and State University

This paper was prepared for delivery at the annual meetings of the American Educational Research Association, New Orleans, April 23-27, 1984. We would like to thank Robert M. Hauser for his very helpful comments on an earlier draft of this paper.

WITHIN-VARIABLE, BETWEEN-OCCASION ERROR COVARIANCES IN MODELS OF EDUCATIONAL ACHIEVEMENT

ABSTRACT

To correct for the effects of measurement error on structural parameter estimates, many researchers are now estimating models of educational achievement with LISREL. In order to estimate such models it is desirable to obtain multiple manifest measures of the latent constructs. Many researchers restrict their models to two manifest measures per latent construct for reasons of economy; but doing so assumes in the absence of external information that all of the covariance between the within-variable measures is reliable covariance. Such an assumption may or may not hold in practice. The present study empirically investigates the extent of within-variable, between-occasion error covariances among variables typically included in models of educational achievement using data from High School and Beyond.

WITHIN-VARIABLE, BETWEEN-OCCASION ERROR COVARIANCES IN MODELS OF EDUCATIONAL ACHIEVEMENT

In the presence of random measurement error in the regressors, ordinary least-squares (OLS) estimates are biased downward in bivariate regressions and often biased in multivariate regressions (Walker and Lev, 1953). The usual response has been to assume the effects of measurement error are probably minor, and consequently to report uncorrected OLS estimates. Others have been less sanguine. Bowles (1972) and Bowles and Gintis (1976), among others, have argued that in models of socioeconomic achievement the biases due to measurement error in background variables are serious, and suggested that social class background is considerably more important than indicated by uncorrected OLS estimates.

Recent methodological advances (Joreskog, 1966, 1967, 1969; Joreskog and Sorbom, 1981) permit the empirical estimation of the measurement properties of manifest variables, and allow the estimation of structural parameter estimates corrected for measurement error. Initial applications (e.g., Bielby, Hauser, and Featherman, 1977) of this new procedure (LISREL) to the analysis of models of socioeconomic achievement indicated that the biasing effects of measurement error among social background variables were neither pervasive nor very large.

In brief, the strategy followed by Bielby, et al. (1977), and several other applications that followed (e.g., Bielby and Hauser, 1977; Corcoran, 1980; Hauser, Tsai, and Sewell, 1983; Mare and Mason, 1980;

Wolfle, 1983) has been either to measure retrospective reports of parental status at two points in time or to use multiple respondents, and to use the consistency of such reports as indicators of their reliabilities. This approach generally shows that parental status is reported with reliabilities ranging from .75 to .90.

Several people (Broom, et al., 1978; Heyns, 1978; Jencks, et al., 1979); however, have pointed out that multiple manifest responses may be consistent yet inaccurate. Conventional methods of measuring reliability cannot detect errors of this type. The problem develops from the necessary practice of specifying that the error covariance of within-variable manifest variables is fixed, usually at zero. To do otherwise would specify an underidentified model, which would have no unique solution. The consequence, however, is that all of the covariance between the within-variable measures is estimated to be reliable covariance. Such a specification is unwarranted if response errors are consistent over time. Respondents who overestimate or underestimate a socioeconomic measure on one occasion may repeat their mistake on another occasion. With only two manifest measures, however, these correlated response errors are included in the estimate of reliable covariance, and hence tend to overestimate the reliability of social background variables.

The typical response to the possible presence of within-variable, between-occasion correlated response errors has been to ignore the problem, no doubt in the hope that the consequences were minor and insignificant. Some authors, however, have tried to circumvent the

identification problem by using external information to fix error covariances to some specified value other than zero. For example, Bielby, et al. (1977) fixed the within-variable, between-occasion error correlations among their socioeconomic variables to the arbitrary value of .5; in a more recent application, Hauser, et al. (1983) estimated error correlations in one sample, then used those estimates to fix the within-variable error correlations for reports of father's education and mother's education at .15 in their analyzed sample.

The present study empirically investigates the extent of within-variable, between-occasion error covariances in models of educational achievement. The data used for this investigation were taken from "High School and Beyond" (HSB), a longitudinal study of U.S. high school sophomores and seniors, sponsored by the National Center for Education Statistics.[1] In particular, these analyses were based on a sample of 3367 HSB parents matched to their sophomore high school children. In the 1980 base-year survey, both children and a subsample of parents were asked to report the educational attainment of the mother and father, and the father's occupation. In 1982, the children were surveyed again, and again asked to report on their parents' socioeconomic statuses. The analysis reported here is restricted to 1064 white respondents who possessed complete reports for the nine variables included in the measurement model.

MEASUREMENT MODEL

The basic measurement model can be described by a set of nine equations. The first three equations consider the parent's report and

both of the student's reports of the father's occupation to be dependent upon the father's true occupational status. The next three equations consider the parent's report and both of the student's reports of the father's education to be dependent upon the father's true educational attainment. The final three quations do the same for mother's education. The three true scores were allowed to covary without constraint. Covariances among the response errors were initially set at zero on the assumption that response errors were random, thus providing a base-line model against which to test subsequent alterations to the model. The model was then modified to permit the estimation of within-variable, between-occasion error covariances; this was accomplished by allowing the errors of the student's reports in the base-year survey to covary with their reports in the first follow-up survey for each of the three parental status variables.

When all three parental status characteristics are considered simultaneously another possible source of error covariance must be appraised. One may suspect that respondents who overestimate or underestimate one parental status measure are also likely to do the same in subsequent reports. One may also suspect that students who are unsure about one parent's education may reconcile their uncertainty by substituting known information about one parent for unknown information about the other (see, e.g., Wolfle and Robertshaw, 1983). In the present analysis, these between-variable, within-occasion error covariances were allowed to be free after first estimating the within-variable, between-occasion error covariances.

The resulting measurement model is shown in Figure 1. The straight arrows represent assumed causal relationships among the latent, true scores (shown in ellipses) and their manifest indicators. Curved, double-headed arrows represent correlations; the correlations among response errors were added to the model during the analysis. For each latent variable, one manifest variable was set equal to the underlying factor plus an error term. In each case, this variable was the parental report of the respective status characteristic, and consequently gives to the underlying factor the same scale of measurement as the parent's report. The two student's reports were set equal to a weighted function of the underlying factor plus an error term.

Insert Figure 1 About Here

To test whether the within-variable, between-occasion error covariances, followed by the between-variable, within-occasion error covariances, provided a significant improvement in the fit of the model to the data, a hierarchical series of models was estimated. The strategy available for selecting the best-fitting model consists, first, of estimating a constrained model (e.g., error covariances set at zero), and then estimating a less constrained model in which one or more formerly fixed parameters are set free. The statistical test consists of assessing the statistical significance of the improvement in fit going from the more constrained to the less constrained model. If the more constrained model fits the data as well as the less constrained model within sampling error

limits, then one may conclude that the constraints do not seriously erode the fit of the model. If the less constrained model yields a significantly better fit, then the less constrained model is probably to be preferred.

RESULTS

Maximum-likelihood estimates of parameters implied by the model shown in Figure 1 were obtained by using LISREL (Joreskog and Sorbom, 1981). The steps followed in the analysis were dictated by the substantive questions. First, a model was estimated in which all error covariances were specified to be zero. Second, a model was estimated in which within-variable, between-occasion error covariances were allowed to be free. Third, several possible between-variable, within-occasion error covariances were tested for statistical significance.

Table 1 presents goodness-of-fit statistics for these several models. Model A of Table 1 is a model with random errors in which all error covariances were specified to be zero. The likelihood ratio chi-square value for Model A was 85.70 with 24 degrees of freedom, indicating that the model as specified did not adequately reproduce the observed covariance matrix.

Insert Table 1 About Here

Model B is a modification of the initial model, in which three within-variable, between-occasion error covariances were set free. These error covariances are between response errors of the student's reports of father's occupation, father's education, and mother's education

in the 1980 base-year survey and the 1982 first follow-up survey. The difference in chi-squares between Model A and Model B is itself distributed as chi-square. This difference was .88 with 3 degrees of freedom. A value this small indicates that none of the three error covariances were statistically significant, a result confirmed by examination of the parameter estimates themselves in relation to their standard errors. It is obvious that HSB respondents reported parental status characteristics with error, for the error variances were all statistically significant. These results suggest, however, that these response errors were not systematically related over time. Apparently the mistakes students made in reporting their parents' status characteristics at one point in time were unrelated to the errors they made in a subsequent survey.

Although the within-variable, between-occasion error covariances may be considered zero within sampling error limits, another possible source of covariance among error terms exists. These are within-occasion, between-variable error covariances, and may exist because respondents within a given survey panel reported status variables with greater consistency than warranted in reality. Previous investigations have found such nonzero error covariances among education variables. Bielby, et al. (1977) reported correlated errors between father's education and respondent's education for blacks. Mare and Mason (1980) reported correlated errors between mother's and father's education for sixth and ninth graders. Wolfle and Robertshaw (1983) found correlated errors among reports of parental education for a sample of 1972 high

school seniors, a result duplicated in Wolfle's (1983) study of 1980 high school seniors. And Hauser, et al. (1983) found such within-occasion correlated response errors not only when the respondents were high school seniors but also when they were resurveyed 18 years later.

To test whether the within-occasion, between-variable error covariances among the students' reports of their parents' status were significantly different from zero, Model C was estimated. Model C included four error covariance terms to be estimated by LISREL: between father's occupation and father's education in the base-year and follow-up surveys, and between father's education and mother's education in both surveys. The chi-square value for Model C as shown in Table 1 was 38.91 with 20 degrees of freedom. The decrease in chi-square from that of Model A was 46.79 with 4 degrees of freedom. A value this large indicates that the error covariances are significantly different from zero.

An examination of the LISREL estimates for Model C indicated that only the error covariances for father's and mother's education were statistically significant. Apparently, when students reported their father's occupation and education they did not report these with greater consistency than warranted by the association of the true scores for these variables. Given these results, Model D was estimated in which the error covariances for father's occupation and father's education were set at zero. Comparing Model D to Model C in Table 1 indicates that these restrictions did not affect the fit of the model within sampling error limits.

A further examination of the estimates in Model D revealed that one final change in the model was warranted. The modification indices in LISREL indicate which fixed parameters in the model, if set free, will provide improvements in the fit of the model to the data. These indicated that the error covariance between the parental reports of father's and mother's education should be set free. Accordingly, Model E was estimated in which all three between-variable, within-occasion error covariances for father's and mother's education were allowed to be free.

As shown in Table 1, the chi-square value for Model E was 28.24 with 21 degrees of freedom, and indicated that the model adequately fit the data and was an improvement over that of Model D. All of the free estimates in the model were statistically significant, and none of the modification indices indicated that any fixed parameters, if set free, would significantly improve the fit of the model to the data.

The parameter estimates for Model E are shown in Table 2, and in standardized form in Figure 1. These values indicate, first, that parental educational attainment was reported more accurately than father's occupational status. This finding conflicts with Mare and Mason (1980), who found just the opposite, but agrees with the results of Wolfle and Robertshaw's (1983) study of measurement error in 1972 high school seniors' reports of parental status characteristics. Second, the estimates shown in Table 2 indicate that the reliability estimates for students are higher in value than those of parents, differences which are statistically significant. At first blush this seems implausible. We

are inclined to attribute this anomaly to the fact that the students were responding to a written questionnaire while the parents were responding to a telephone survey, which may have introduced additional sources of reporting and coding errors into the data. Finally, while the within-variable, between-occasion error covariances have been found to be zero for these data, between-variable, within-occasion error covariances for father's and mother's education were found to be statistically significant. As shown in Figure 1, the correlations among these error terms range from .14 to .20, indicating that both parents and students reported parental educational levels with greater consistency than warranted by the association of their true educational attainments.

Insert Table 2 About Here

CONCLUSION

The process of educational and occupational achievement has become a major line of social science research. Lately, however, the findings in this line have been called into question because they may have depended on variables subject to large response errors (e.g., Bowles, 1972). Early applications of LISREL-type models that compensated for the effects of measurement error on structural parameter estimates indicated that the concern was unwarranted (Bielby, et al., 1977). Subsequently, researchers became concerned that models which depended on dual reports of background characteristics overestimated the reliability of these variables by including correlated errors among estimates of reliable covariance (Broom, et al., 1978; Jencks, et al.,

1979). Using data from a recent nationwide survey of high school sophomores in 1980, this study investigated the extent of these within-variable, between-occasion error covariances. In short, we found no evidence to indicate the existence of such error covariances in the reports of high school sophomores resurveyed two years later. Like previous studies in this area, our data yield estimates of reliability that are in the range of .7 to .9 for students' reports of parental status characteristics. Given the findings presented here, we do not believe further concern is warranted about whether these reliability estimates are artificially inflated by correlated errors.

The presence of between-variable, within-occasion error covariances is another matter. Massagli and Hauser (1983) noted in their review of a number of studies that "none of the studies has provided substantial evidence of nonrandom measurement error." Our findings indicate that models which include measures of father's occupational status and father's education are generally free of nonrandom error. Nevertheless, we have found evidence that when respondents are asked to report both parents' educational attainment they do so with correlated errors of measurement. Our conjecture is that respondents reconcile uncertainty about one parent's educational attainment by making reference to perceptions of the other parent's education. The significance of these correlated response errors notwithstanding, their influence on the estimated correlation between father's and mother's true educations is modest; with correlated response errors (Model E) the estimated correlation was .621; without correlated response errors (Model A) the estimated correlation was .627.

In conclusion, we find little evidence to support the claim that reliability estimates for social background variables are inflated due to correlated errors of measurement. We tested for such correlated errors among reports of father's occupational status, father's education, and mother's education, and found none. We did find, however, that when respondents are asked to report parental educational attainments, they tend to do so with greater consistency across parental reports than warranted by the true association of parents' education. Given these findings, coupled with the presence of simple unreliability, we are led more than ever to prefer analyses of educational and occupational achievement based on models with well-defined measurement and error structures.

FOOTNOTES

1. The exact questions used in the original survey are available in the data file user's manual (National Opinion Research Center, 1983), but are summarized here. Of the variables included in this analysis, the HSB sophomores were first asked to categorize the job most recently held by their father. They were asked to choose one of seventeen categories (clerical, craftsmen, farmer, etc.); these responses were then recoded to their equivalent Duncan (1961) socioeconomic index scores as given in Levinsohn, et al. (1978, Appendix O, p. 11). The sophomores were next asked to indicate the highest level of education completed by their father. A similar question was asked about their mother's education. These responses were then recoded to match the categories used in the parent's survey; the resulting scale ranged from 1 to 8, representing categories from less than high school (=1) to the receipt of a Ph.D., M.D., or other advanced degree (=8).

After the collection of base-year data from the high school students, 3367 parents of the HSB sophomores were contacted and additional data collected, which concentrated primarily on the parents' plans for financing their children's higher education. Included in the questionnaire, however, were questions dealing with parental socioeconomic characteristics. In about 60 percent of the cases, it was the student's mother who completed the questionnaire, while the student's father completed the questionnaire in the remaining cases (students who had some other adult complete the questionnaire, such as an aunt or grandfather, were excluded from these analyses). Parents completing the questionnaire were asked to report their occupation, their spouse's occupation, their education, and their spouse's education. These were recoded as appropriate to obtain a report of the father's education (as reported either by himself or his spouse), and mother's education. These were recoded to match equivalently the scale used by students to report their parents' education. The occupation question used in the parent's survey was coded according to the U.S. Census Bureau's detailed occupation code. In order to match these responses with those of the high school sophomores, the detailed occupation codes were collapsed into the identical categories used by the students, and assigned the same Duncan (1961) SEI scores.

In 1982, the HSB sophomore cohort, now high school seniors, was resurveyed. They were once again asked to report their parents' statuses, and their responses were treated exactly as described above for the base-year data.

The correlations among these nine variables, plus their means and standard deviations, are shown in Appendix Table A.

Insert Table A About Here

REFERENCES

Bielby, W.T., & Hauser, R.M. (1977). Response error in earnings functions for nonblack males. Sociological Methods & Research, 6, 241-280.

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

Bowles, S. (1972). Schooling and inequality from generation to generation. Journal of Political Economy, 80, S219-S251.

Bowles, S., & Gintis, H. (1976). Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life. New York: Basic Books.

Broom, L., Jones, F.L., McDonnell, P., & Duncan-Jones, P. (1978). Is it true what they say about daddy? American Journal of Sociology, 84, 417-426.

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

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

Hauser, R.M., Tsai, S.L., & Sewell, W.H. (1983). A model of stratification with response error in social and psychological variables. Sociology of Education, 56, 20-46.

Heyns, B. (1978). Review essay: Schooling in capitalist America. American Journal of Sociology, 83, 999-1006.

Jencks, C., Bartlett, S., Corcoran, M., Crouse, J., Eaglesfield, D., Jackson, G., McClelland, K., Mueser, P., Olneck, M., Schwartz, J., Ward, S., & Williams, J. (1979). Who Gets Ahead? The Determinants of Economic Success in America. New York: Basic Books.

Joreskog, K.G. (1966). UMFLA: A Computer Program for Unrestricted Maximum Likelihood Factor Analysis. Research Memorandum 66-20. Princeton, NJ: Educational Testing Service.

Joreskog, K.G. (1967). Some contributions to maximum likelihood factor analysis. Psychometrika, 31, 165-178.

Joreskog, K.G. (1969). A general approach to confirmatory maximum likelihood factor analysis. Psychometrika, 34, 183-202.

Joreskog, K.G., & Sorbom, D. (1981). LISREL: Analysis of Linear Structural Relationships by the Method of Maximum Likelihood. Chicago: National Educational Resources.

Levinsohn, J.R., Henderson, L.B., Riccobono, J.A., & Moore, R.P. (1978). National Longitudinal Study: Base Year, First, Second and Third Follow-Up Data File Users Manual. Washington, D.C.: National Center for Education Statistics.

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

Massagli, M.P., & Hauser, R.M. (1983). Response variability in self- and proxy reports of paternal and filial socioeconomic characteristics. American Journal of Sociology, 89, 420-431.

National Opinion Research Center. (1983). High School and Beyond 1980 Sophomore Cohort First Follow-Up (1982) Data File User's Manual. Washington, D.C.: National Center for Education Statistics.

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

Wolfle, L.M. (1983, April). Postsecondary Educational Attainment among Whites and Blacks. Paper presented at the meeting of the American Educational Research Association, Montreal, Quebec.

Wolfle, L.M., & Robertshaw, D. (1983). Racial differences in measurement error in educational achievement models. Journal of Educational Measurement, 20, 39-49.

Table 1: Goodness-of-Fit Statistics for Measurement Models of Parental Status Characteristics, White High School Sophomores in 1980 (N = 1064 listwise present)

Model	L^2	d.f.	Prob.	L^2 /d.f.
A. No error covariances	85.70	24	.000	3.57
B. Within-variable, between-occasion/error covariances	84.82	21	.000	4.04
B vs. A	.88	3	.830	.29
C. Between-variable, within-occasion error covariances for students' reports of parental status	38.91	20	.007	1.95
C vs. A	46.79	4	.000	11.70
D. Between-variable, within-occasion error covariances for students' reports of parental education	39.08	22	.014	1.78
C vs. D	.17	2	.918	.08
E. Between-variable, within-occasion error covariances for parents' and students' reports of parental education	28.24	21	.133	1.34
E vs. D	10.84	1	.001	10.84

Table 2. Model E Parameter Estimates
(standard errors in parentheses)

Latent Characteristic	Manifest Report	Lambda Coefficients	Error Variance	Reliability
Father's Occupation	Parent	1.00*	172.06 (9.90)	.66
	Stud. 1	1.017 (.034)	144.23 (9.03)	.71
	Stud. 2	1.026 (.033)	118.64 (8.27)	.75
Father's Education	Parent	1.00*	.552 (.032)	.88
	Stud. 1	1.061 (.016)	.421 (.030)	.92
	Stud. 2	1.062 (.015)	.380 (.029)	.93
Mother's Education	Parent	1.00*	.404 (.025)	.86
	Stud. 1	1.080 (.019)	.419 (.027)	.88
	Stud. 2	1.097 (.018)	.301 (.025)	.91

True Score Covariances

	1.	2.	3.
1. Father's Occup.	334.71 (21.67)		
2. Father's Educ.	27.45 (1.64)	4.23 (.21)	
3. Mother's Educ.	13.44 (1.11)	2.05 (.13)	2.59 (.13)

Error Covariances

1. Fath. Educ. Parent and Moth. Educ. Parent	.064 (.020)
2. Fath. Edud. Stud. 1 and Moth. Educ. Stud. 1	.066 (.020)
3. Fath. Educ. Stud. 2 and Moth. Educ. Stud. 2	.068 (.018)

*Fixed value.

Table A. Correlations, Means, and Standard Deviations Among Variables in Measurement Model of Parental Status Characteristics, White High School Sophomores in 1980
(N = 1064 listwise present)

Respondent Variable		1	2	3	4	5	6	7	8	9
1. Parent	Fa.Occ.	1.00								
2. Stud. 1	Fa.Occ.	.675	1.00							
3. Stud. 2	Fa.Occ.	.710	.726	1.00						
4. Parent	Fa.Educ.	.567	.597	.592	1.00					
5. Stud. 1	Fa.Educ.	.565	.597	.590	.901	1.00				
6. Stud. 2	Fa.Educ.	.565	.600	.597	.905	.923	1.00			
7. Parent	Ma.Educ.	.371	.395	.365	.559	.561	.553	1.00		
8. Stud. 1	Ma.Educ.	.356	.383	.334	.554	.586	.564	.870	1.00	
9. Stud. 2	Ma.Educ.	.348	.386	.342	.546	.569	.576	.888	.895	1.00
Means		45.94	43.56	45.48	3.744	3.644	3.659	3.128	3.156	3.189
Standard Dev.		22.51	22.15	21.71	2.187	2.279	2.266	1.729	1.856	1.846

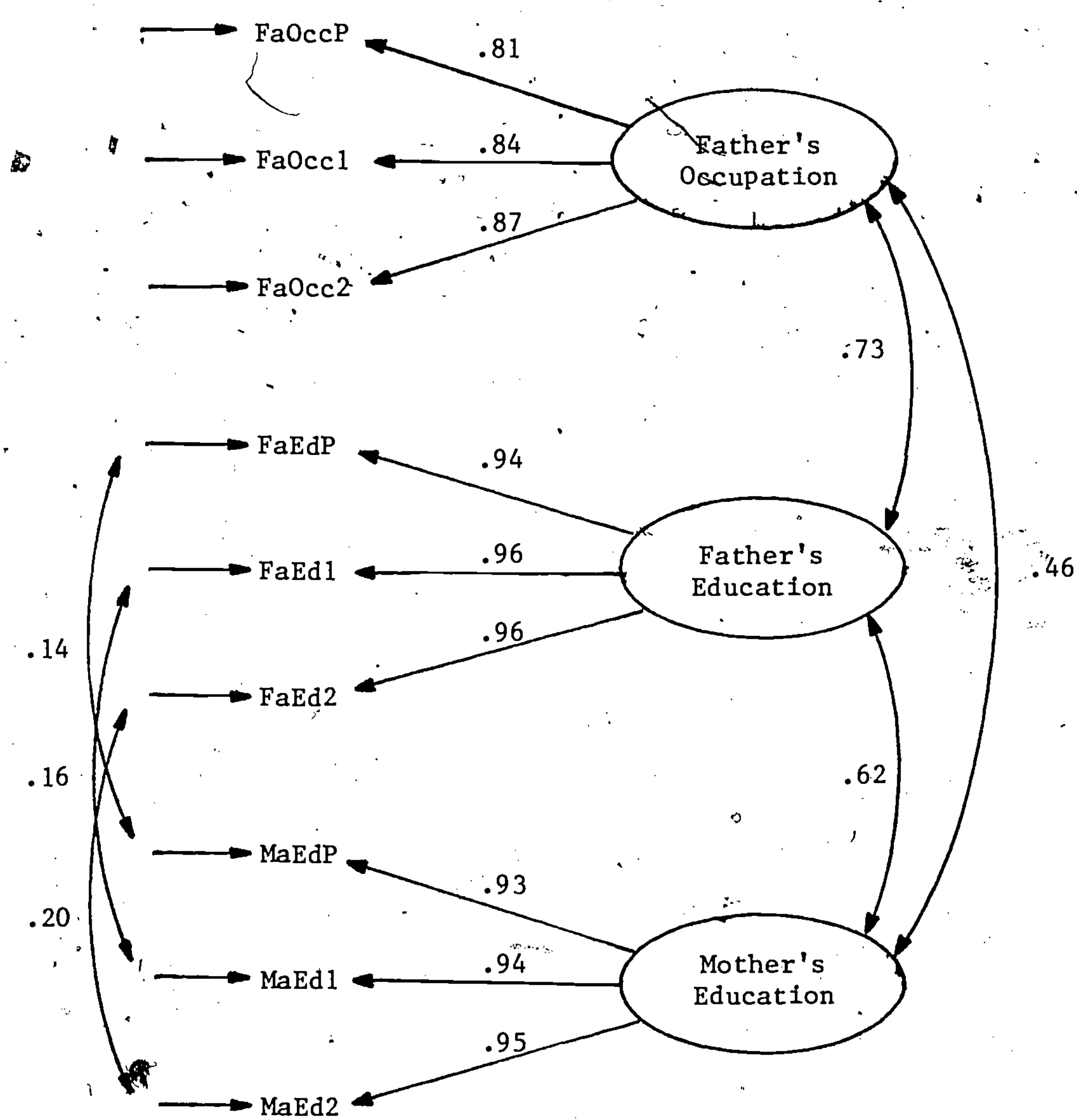

Figure 1. Measurement Model with Correlated Response Errors