DOCUMENT RESUME

ED 053 156 TM 000 665

AUTHOR McGaw, Barry; Joreskog, Karl G.
TITLE Factorial Invariance of Ability Measures in Groups Differing in Intelligence and Socioeconomic Status.
INSTITUTION Educational Testing Service, Princeton, N.J.
REPORT NO RB-70-63
PUB DATE Nov 70
NOTE 31p.

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS *Ability Identification, Achievement Tests, Analysis of Covariance, Analysis of Variance, Aptitude Tests, Comparative Analysis, *Factor Analysis, Goodness of Fit, *Groups, High School Students, Intelligence Differences, *Intelligence Level, Mathematical Models, Scores, *Socioeconomic Status
IDENTIFIERS Project TALENT

ABSTRACT

This study attempts to determine whether the pattern of human abilities varies or remains constant over a range of ability levels. Scores on 12 aptitude and achievement tests for 11,743 subjects, subdivided into four groups according to intelligence and socioeconomic status, were used. A technique, developed by Joreskog, for simultaneously factor analyzing data from several populations was used to determine whether there was factorial invariance over the four groups. A model, in which the same factor pattern (matrix of factor loadings) was assumed to hold for the four groups, was fitted to the data. Goodness of fit indices suggested the model fitted satisfactorily. Differences in the factor dispersion matrices and mean factor scores for the subpopulations were then examined and discussed. Statistical data is included in the accompanying tables. (Author/LR)

RB-70-63

FACTORIAL INVARIANCE OF ABILITY MEASURES IN GROUPS DIFFERING IN INTELLIGENCE AND SOCIOECONOMIC STATUS

Barry McGaw
University of Illinois

and

Karl G. Jöreskog
Educational Testing Service

This Bulletin is a draft for interoffice circulation. Corrections and suggestions for revision are solicited. The Bulletin should not be cited as a reference without the specific permission of the authors. It is automatically superseded upon formal publication of the material.

Educational Testing Service
Princeton, New Jersey
November 1970

FACTORIAL INVARIANCE OF ABILITY MEASURES IN GROUPS DIFFERING IN INTELLIGENCE AND SOCIOECONOMIC STATUS

Barry McGaw

University of Illinois

and

Karl G. Jöreskog

Educational Testing Service

ABSTRACT

Scores on 12 aptitude and achievement tests for 11,743 subjects, subdivided into four groups according to intelligence and socioeconomic status, were used. A technique, developed by Jöreskog, for simultaneously factor analyzing data from several populations was used to determine whether there was factorial invariance over the four groups.

A model, in which the same factor pattern (matrix of factor loadings) was assumed to hold for the four groups, was fitted to the data. Goodness of fit indices suggested the model fitted satisfactorily. Differences in the factor dispersion matrices and mean factor scores for the subpopulations were then examined and discussed.

FACTORIAL INVARIANCE OF ABILITY MEASURES IN GROUPS DIFFERING IN INTELLIGENCE AND SOCIOECONOMIC STATUS[1]

Barry McGaw

University of Illinois

and

Karl G. Jöreskog

Educational Testing Service

An important issue in the assessment of human abilities is whether, over a range of ability levels (or levels of other variables), the pattern of abilities remains constant. For example, over a range of levels of socio-economic status, or intelligence, does the pattern of abilities vary or remain essentially the same?

In this study a sample of 11,743 high school subjects was divided into four groups formed as the four combinations of high and low intelligence and high and low socioeconomic status. Ability measures were obtained for the subjects on 21 different variables. The data used were collected in the Project TALENT study.[2] The chief concern of the present study was to determine whether the same factor pattern could be obtained for the four groups.

Factorial invariance has been used in a number of studies to assess the similarity of different groups. Johnson (1969) found factorial invariance of educational abilities and aptitudes with data from subjects in Rhodesia

---

[1]This study was completed while the first author was working with the second author during the Summer Program for Graduate Students at Educational Testing Service during the summer of 1970.

[2]The data used were kindly made available by Dr. L. G. Humphreys, University of Illinois.

and Zambia. Rock and Freeberg (1969) used a similar technique to examine the factor stability of a student biographical information blank administered at different grade levels. The technique used in both of these studies was that suggested by Meredith (1964a,b). Meredith's procedure involves an attempted rotation to factorial invariance. For each group an orthogonal factor solution is first obtained, then an attempt is made to rotate these to a common orthogonal solution. If this can be achieved there remains some freedom for further transformation to a more readily interpretable oblique solution.

In the present study, however, a general procedure, developed by Jöreskog (1971) for simultaneous factor analysis in several populations, was used. This procedure is similar to that of Lawley and Maxwell (1963, Chapter 8) but is more general. It provides for testing of the hypothesis of factorial invariance. Rather than attempting to rotate independent solutions to a common solution, the model to be tested may be specified a priori, and the computer program then estimates the model using the data from all the populations simultaneously by the maximum likelihood method, and provides a measure of goodness of fit.

The essential difference between these two approaches is that Meredith's involves a search to determine whether there is a set of rotations which can transform independent solutions to a common solution while Jöreskog's involves the direct fitting of an hypothesized model of factorial invariance to the data.

Subjects and Data

The subsamples obtained by the subdivision of the sample described above were low IQ-low SES, $N_1 = 4491$; low IQ-high SES, $N_2 = 1336$;

high IQ-low SES, $N_3 = 939$; high IQ-high SES, $N_4 = 4977$.

The variables used in the Project TALENT study are described in detail elsewhere (e.g. Flanagan et al., 1962). The 21 measures used initially in this study were the subscores from the Information Test (vocabulary, part I, part II) and all of the remaining tests except Reading Comprehension and Abstract Reasoning.

Data Analysis

Dispersion (variance-covariance) matrices S_1 , S_2 , S_3 , and S_4 were computed for the four subsamples, respectively. Because differences in the scales of different tests are arbitrary, it is usual, for the factor analysis of data from a single population, to scale the dispersion matrix to a correlation matrix. In the present study, however, scaling each of the four dispersion matrices, S_1 , S_2 , S_3 , and S_4 , to correlation matrices would remove important differences among the groups on the tests. In principle, the analysis should be performed on the dispersion matrices. These matrices may be rescaled, provided they are kept on a common metric. A convenient rescaling is one in which a weighted average of the rescaled dispersion matrices is a correlation matrix.

Firstly, a pooled dispersion matrix S was calculated as

$$S = \frac{\sum_{g=1}^{4} (N_g - 1)S_g}{\sum_{g=1}^{4} (N_g - 1)} .$$

Then, from the pooled dispersion matrix, a pooled correlation matrix R was calculated as

$$R = DSD$$

where $D = (\text{Diag } S)^{-1/2}$. Finally the original dispersion matrices S_g , $g = 1,2,3,4$, were rescaled to

$$S^*_g = DS_gD \quad .$$

The S^*_g matrices (Table 1) are dispersion matrices, but under this rescaling their weighted average is a correlation matrix. In the analyses which follow the pooled correlation matrix and the rescaled dispersion matrices will be used.

(Table 1 about here)

A preliminary test of the equality of the population dispersion matrices, from which these four sample dispersion matrices were obtained, using a test developed by Box (1949) revealed significant differences ($F_{\infty,\infty} = 38.8$).

Preliminary Factor Analysis

A maximum likelihood factor analysis followed by a varimax rotation was run on the pooled correlation matrix, extracting successively zero through eight common factors. The purpose of this analysis was firstly to determine the number of common factors which appeared to give a satisfactory solution for the composite data and secondly to provide a basis for selecting a reduced number of tests for subsequent analysis. This selection of tests was necessitated by the limitations of the computer program to be used in the simultaneous factor analysis of the four rescaled dispersion matrices.

This initial factor analysis was done using Jöreskog's (1967a,b) unrestricted maximum likelihood factor analysis (UMLFA) which gives a large sample χ^2 test of goodness of fit of the factor model. When the sample size is very large, however, as in this case with $N = 11743$, the χ^2 statistic is very sensitive to minor departures from the model. With the present data a χ^2 of 403 with 70 degrees of freedom was obtained for a solution with 8 common factors, indicating that some larger number of common factors was required. This difficulty with the chi square test of goodness of fit has been discussed by Cochran (1952) and Gulliksen and Tukey (1958). The latter authors, in considering the Law of Categorical Judgments, used a variance components analysis in order to determine whether the variance accounted for by the theory is large or small in relation to the total variance in the data. Tucker and Lewis (1970) have developed a similar approach for factor analysis. They have suggested a reliability index ρ which may be estimated as

$$\hat{\rho} = \frac{M_0 - M_k}{M_0 - 1}$$

where $M_0 = \chi_0^2/df_0$ and $M_k = \chi_k^2/df_k$, the χ^2's and degrees of freedom being those obtained for maximum likelihood factor solutions with zero and k common factors. The value of $\hat{\rho}$ provides a measure of goodness of fit for the particular factor model, with k common factors being fitted. This index takes into account not only sampling error, as the χ^2 does, but also discrepancies between the population dispersion matrix and the formal factor model (specification errors). The upper bound on this index, obtained when the model perfectly fits the data, is unity.

(Table 2 about here)

For the models, with zero through eight common factors, tested on the pooled correlation matrix of 21 variables, the indices of goodness of fit are shown in Table 2. On the basis of these analyses it was decided that a four factor solution appeared reasonable. Twelve of the original variables were then selected so that each of the factors would be identified by at least three variables. An unrestricted maximum likelihood factor analysis of the pooled correlation matrix for these twelve variables yielded the goodness of fit indices shown in Table 3, for solutions with zero through four common factors.

(Table 3 about here)

Restricted Maximum Likelihood Factor Analysis

Restricted maximum likelihood factor analysis is a method described by Jöreskog and Lawley (1968) in which some parameters of the factor model may be fixed, *a priori*. With this technique the appropriateness of variously specified factor models can be investigated.

The basic factor analysis model is

$$x = \Lambda f + z$$

where x is a vector of order p of observed test scores, f is a vector of order $k < p$ of latent common factor scores, z is a vector of order p of unique scores, and Λ is a $p \times k$ matrix of factor loadings (the factor pattern). It is assumed that f and z are independent random vectors with $E(f) = 0$, $E(z) = 0$, $E(ff') = \Phi$, and $E(zz') = \psi^2$, a diagonal matrix. From these assumptions the dispersion matrix Σ of x can be shown to be

$$\Sigma = \Lambda\Phi\Lambda' + \psi^2$$

where Φ is the factor dispersion matrix and ψ^2 the diagonal matrix of unique variances.

(Table 4 about here)

The rotated varimax solution from the unrestricted factor analysis of the pooled correlation matrix for the 12 selected variables is shown in Table 4. This solution was the starting point for several trial solutions with correlated factors obtained with the ACOVS program (Jöreskog et al., 1970). In order to achieve a unique solution with k common factors, it is necessary to have at least k^2 fixed elements in Λ and Φ (Jöreskog, 1969) of which at least $k - 1$ must be in each column of Λ. The further k restrictions may be imposed in Λ or Φ, for example, by fixing the diagonal elements of Φ to unity and thus making Φ a correlation matrix. As a first step this minimum number of restrictions was imposed by constraining Φ to be a correlation matrix and fixing three zeros in each column of Λ. Finally a relatively well fitting model was achieved with as many zero elements in Λ as seemed theoretically meaningful and indicated by the data. This solution is shown in Tables 5a-b.

(Tables 5a-b about here)

The goodness of fit test yielded a χ^2 of 852.9 with 41 degrees of freedom. The unrestricted solution with four factors yielded a χ^2 of 202.15 with 24 degrees of freedom. Tucker and Lewis' reliability index provides a more ready basis for comparison. For the unrestricted solution $\hat{\rho}$ was .987 and for the restricted solution .965. An unrestricted model, in general, provides a better fit than a restricted model, but the restricted model, with fewer parameters to be estimated, is more parsimonious. In the present case the restricted model appears to fit satisfactorily.

The factors are readily interpretable. Factor I is a general knowledge factor. Although in the mechanical reasoning test "every item can be answered without training in physics, and without experience in woodworking or other crafts, or in working with motors" past training and experience could well have a facilitative effect (see Flanagan et al., 1962, p. 109).

Factor II is a verbal mechanics factor. The loading of the clerical checking test on this factor is interesting. On this test subjects are required to determine quickly and accurately whether pairs of names are the same, this restriction of items to words apparently introducing a purely verbal component in addition to the perceptual speed.

Factor III is a spatial perception factor with, probably, a three-dimensional component since the highest loadings are for those tests which clearly require three-dimensional perception.

Factor IV involves speed in two-dimensional perception. The three tests, Table Reading, Clerical Checking and Object Inspection, were designed to measure speed and accuracy of perception and all are loaded on this factor. The loading of Visualization I (two-dimensional) indicates the facilitative effect of rapid two-dimensional perception on the tasks required in that particular test. The loading of the Object Inspection test on Factor III probably is a result of the use of three dimensions in the object diagrams. The differences can be perceived without three-dimensional perception but they obviously would be perceived more readily if the full dimensionality of the object were seen.

Simultaneous Factor Analysis in the Four Populations

Jöreskog has developed a technique for simultaneously factor analyzing data from several populations. The procedure is described in detail elsewhere (Jöreskog, 1971). The model for each population is

$$x_g = \mu + \Lambda f_g + z_g$$

where x_g is a vector of order p of observed test scores from the g^{th} population, μ is the vector of order p of overall mean test scores from all the populations involved, f_g is a vector of order $k < p$ of latent common factor scores, z_g is a vector of order p of unique scores, and Λ is a $p \times k$ matrix of factor loadings (factor pattern). Jöreskog's general model and the computer program SIFASP (van Thillo and Jöreskog, 1970) will accommodate the more general case where there is a different factor pattern Λ_g for each population, but in the present study we are concerned only with the case where the same factor pattern can be fitted to all populations.

If ν_g is the vector of order k of mean factor scores for the g^{th} population, i.e.,

$$E(f_g) = \nu_g$$

then, without loss of generality, the origin of the factor scores can be fixed so that

$$\Sigma N_g \nu_g = 0 \quad .$$

If it is assumed that the usual factor analysis model holds in each population, with the same Λ, the respective population dispersion matrices can be represented as

$$\Sigma_g = \Lambda \Phi_g \Lambda' + \psi_g^2 \quad .$$

The elements of Λ, Φ_g, and ψ_g^2 $(g = 1,2,\ldots,G)$ are the parameters to be estimated from the data. If, from each population, there is a random sample of $N_g = n_g + 1$ observations of x_g the usual variance-covariance estimates can be found as the matrices S_g. If x_g has a multivariate normal distribution in each population the likelihood function can be set up and numerically maximized, as Jöreskog has shown, to obtain the maximum likelihood estimates of the parameters.

The log-likelihood function for the g^{th} group is given by

$$\log L_g = -\frac{1}{2} n_g \left[\log|\Sigma_g| + tr(S_g \Sigma_g^{-1})\right] \quad .$$

If the samples of subjects are independent, the likelihood function L for all groups is given by

$$L = \prod_{g=1}^{G} L_g$$

and, thus, the log-likelihood function is

$$\log L = \sum_{g=1}^{G} \log L_g \quad .$$

The maximization of the likelihood function is achieved through the minimization, for each population, of the function

$$F_g(\Lambda,\Phi_g,\psi_g) = n_g\left[\log|\Sigma_g| + \mathrm{tr}(S_g\Sigma_g^{-1}) - \log|S_g| - p\right] \quad .$$

For all G groups, the function to be minimized is

$$F(\Lambda,\Phi_1,\ldots,\Phi_G,\psi_1,\ldots,\psi_G) = \sum_{g=1}^{G} F_g \quad .$$

This function F is a function of the free parameters in Λ, the free parameters in the lower halves of the Φ_g matrices, including the diagonals, and the free parameters in the diagonals of the ψ_g^2 matrices.

As it stands this function is unaffected by an arbitrary linear transformation of Λ. If, for any nonsingular $k \times k$ matrix T, Λ is transformed to Λ^* by

$$\Lambda^* = \Lambda T^{-1}$$

and a complementary transformation of the factor axes is performed as

$$\Phi_g^* = T\Phi T' \qquad \text{for } g = 1,2,\ldots,G$$

the function F remains unaltered.

In order for the parameters to be defined uniquely there must be imposed at least k^2 conditions on the Λ and/or all the Φ_g's. These conditions are imposed by fixing elements in advance. The final solution can be obtained in either of two ways. The equivalent of an unrestricted solution may be obtained by imposing some k^2 arbitrary values and subsequently rotating. On the other hand, if a particular model is being tested the model may be specified by fixing at least k^2 particular parameters and obtaining the solution directly. If more than k^2 conditions are imposed the solution will be restricted and, therefore, could not be obtained exactly by the transformation to similarity of unrestricted solutions.

With the data in the present study a restricted solution was obtained. All the elements in Λ which had been fixed at zero in the restricted solution with the pooled correlation matrix (Table 5a) were similarly fixed for the simultaneous solution. A further four restrictions were imposed by requiring that the weighted mean of the Φ_g matrices be a correlation matrix, viz,

$$\text{diag } \Phi = I$$

where $\Phi = \frac{1}{n} \sum_{g=1}^{4} n_g \Phi_g$. These four restrictions were imposed indirectly by fixing the largest nonzero element in each of the four columns of Λ and thus fixing an arbitrary scaling for both Λ and the four Φ matrices. The solution was then rescaled by calculating Φ as shown above and then, with $D = (\text{diag } \Phi)^{-1/2}$ scaling the factor dispersion matrices Φ_g so that

$$\Phi^*_g = D\Phi_g D \qquad \text{for} \quad g = 1,2,3,4$$

and the common factor pattern matrix Λ so that

$$\Lambda^* = \Lambda D^{-1} \quad .$$

There were, therefore, 15 free parameters in Λ , 40 free parameters in the lower halves (including diagonals) of the four Φ_g matrices, and 48 free parameters in the diagonals of the four ψ_g^2 matrices, giving a total of 103 free parameters in the function to be minimized.

To keep the number of iterations in the numerical minimization to a minimum it was important to obtain good initial estimates of the unknown parameters. The Λ obtained from the final analysis of the pooled correlation matrix (Table 5a) was used as the starting value for Λ . Initial estimates of the Φ_g and ψ_g^2 matrices were obtained by performing separate restricted factor analyses for each of the S_g^* matrices with Λ entirely fixed with the values in Table 5a and the Φ_g and ψ_g^2 matrices entirely free. The solutions for these independent analyses were used as starting points for the simultaneous analysis. The scaled solution for the simultaneous analysis is shown in Tables 6a-b.

(Tables 6a-b about here)

The χ^2 measure of goodness of fit was 2038.2 with 209 degrees of freedom. In order to compute Tucker and Lewis' reliability index, χ^2 values were obtained for a zero common factor solution with each of the S_g^* matrices. These χ^2 and their corresponding degrees of freedom are additive and the totals can be compared with that obtained for the simultaneous solution with 4 common factors. These totals were $\chi^2 = 38741$ and 264 degrees of freedom. The reliability index $\hat{\rho}$ for the simultaneous solution was .940.

As a further check on the goodness of fit of the simultaneous solution independent restricted solutions were obtained for each of the S_g^* matrices with the same elements in Λ, Φ_g, and ψ_g^2 free as in the simultaneous solutions. The sums of the χ^2 and degrees of freedom obtained in these four analyses yielded a χ^2 of 1857.0 with 164 degrees of freedom. The index in this case was .929. It appears, therefore, that the model with the common factor pattern Λ fits the data from the four samples reasonably well.

Mean Factor Scores

In the general model used in the simultaneous factor analysis technique described above factor scores are not standardized within each population. The vector of mean factor scores for each population ν_g can be estimated using Lawley and Maxwell's (1963) modification, for the case of correlated factors, of Thomson's regression method for determining factor scores, viz

$$\hat{\nu}_g = \hat{\Phi}_g \hat{\Lambda}' \hat{\Sigma}_g^{-1} (\bar{x}_g - \hat{\mu})$$

where $\hat{\Sigma}_g = \hat{\Lambda} \hat{\Phi}_g \hat{\Lambda}' + \hat{\psi}^2$ is the estimate of the dispersion matrix for population g, $\bar{x}_g$ is the vector of mean test scores for population g and $\hat{\mu}$ is the vector of overall mean test scores for all populations. The estimates of the mean factor scores shown in Table 7 have been scaled so that $\Sigma N_g \hat{\nu}_g = 0$.

(Table 7 about here)

Discussion

The goodness of fit of the model with Λ invariant over the four populations indicates that there is factorial invariance, in the sense that a solution can be obtained for all populations with the same pattern of factor loadings on the same number of common factors. Whether the common factors are the same in the four populations can not be determined without consideration of the Φ_g matrices, in particular the Φ^*_g matrices obtained in the final scaled solution.

In an extreme case complete equality of the Φ^*_g matrices could be established, either by fitting a model with a common Λ and a common Φ, with all differences in the original S^*_g matrices being accounted for in the diagonal ψ^2_g matrices, or by fitting a model with a common Λ and subsequently testing the hypothesis of equality of the Φ^*_g matrices obtained in the solution. A model with a common Λ and a common Φ allows only for differences in variance elements in the original dispersion matrices and, where these dispersion matrices are demonstrably different as in the present case, there is no point in testing such a model. It is necessary, however, to test for equality of the Φ^*_g matrices obtained in the solution with the common Λ.

The hypothesis of equality of the four factor dispersion matrices Φ^*_g, $g = 1,2,3,4$, shown in Table 6b was tested using the procedure, referred to previously, developed by Box. This test yielded $F_{(30,\infty)} = 89.6$, showing the Φ^*_g matrices to be significantly different. Examination of the matrices reveals the differences to be in both variances and covariances.

Whether, with a common Λ but significantly different Φ_g matrices, the four factors can be given the same substantive interpretation in each population is an important question. It should be noted that there is no mathematical basis for the inference of identity of common factors across populations, even in the case where a common Λ and Φ can be fitted to all populations. It is clearly possible, with different populations, that identical dispersion matrices could be obtained from different test batteries (with the same number of tests in each) applied to the populations. In such a case mathematically identical factor solutions could be obtained.

Where the same test battery is used and the populations are, in fact, subpopulations of some more inclusive population, the inference of identical factors seems reasonable if the Λ and Φ matrices are the same for all subpopulations. While it is probably still appropriate, because of the common Λ, to give the same substantive interpretation of the factors when the Φ_g matrices differ, the differences in the Φ_g matrices reveal differences in the interrelations of the factors in the different subpopulations.

To facilitate a comparison of the interrelations, each of the factor dispersion matrices has been rescaled to a correlation matrix (Table 8). From this table it can be seen that, in the two subpopulations with low SES, Factor IV is more highly correlated with each of the other factors than in the subpopulations with high SES. That is, it appears that perceptual speed is more clearly an independent factor for high SES groups. It is confounded most with the spatial and verbal mechanics factors in the low SES groups. In the two high IQ subpopulations Factors I and II are more highly correlated than in the low IQ groups, with the confounding more marked in the high IQ-low SES population.

(Table 8 about here)

The only other notable difference among the subpopulations is in the correlation between Factors I and III. This correlation is relatively high for the high IQ-high SES subpopulation. Differences in the estimated mean factor scores on the spatial factor indicate that IQ is a more potent variable than SES on this factor. The higher correlation between Factors I and III when SES is higher suggests that the influence of SES on performance on the spatial factor might be mediated through greater general knowledge.

Interpretation of differences in variances on the factors (Table 6b) can be facilitated by simultaneous consideration of differences in mean factor scores (Table 7). The estimates of the mean factor scores given in Table 7 are quite stable, because of the large sample sizes involved. Confidence intervals can be established around each mean since estimates of the variances of the subpopulations on each of the factors are available. The standard error of the estimate of the mean of subpopulation g on factor i is given by

$$S_{\nu_g} = (\phi_{gii}/N_g)^{1/2}$$

where ϕ_{gii} is the i^{th} diagonal element in the factor dispersion matrix Φ^*_g .

Confidence intervals (99%) were established for each of the estimated means shown in Table 7 and, except in the case of the third and fourth subpopulations on the fourth factor, where the estimated means were equal, none of the intervals on any of the factors overlapped. Thus there are real differences in mean factor scores for the groups. This, in itself, is not surprising but what is more interesting are the relationships between these differences and differences in the corresponding variances.

In the following discussion differences in mean factor scores between levels of IQ and SES are compared. Unfortunately, information about the means and variances of the subpopulations on the composite IQ and SES criteria were not available. If, for example, there were greater differences in mean SES at low IQ level than high IQ some of the conclusions that follow might need to be modified. The discussion is based on the presumption that differences between means on one variable (IQ or SES) are the same at both levels of the other variable.

On the general knowledge factor, subpopulation 1 (low IQ-low SES) is well below the other three subpopulations in mean factor score. The gap between low and high SES groups is much wider for low IQ than high IQ. The variances of all but subpopulation 2 (low IQ-high SES) are virtually the same, however. For this subpopulation the dispersion of factor scores is much wider (1.502) suggesting that the interaction of IQ and SES is different at different levels of IQ, the facilitative effect of high SES being much less uniform at low IQ levels.

The IQ variable, as might be expected, appears to be more strongly related than SES to performance on the verbal factor (Table 7), but the facilitative effect of high SES is greater for low IQ (or alternatively, the debilitative effect of low SES is greater for high IQ). The factor scores of all subpopulations are similarly dispersed.

A similar pattern was found on the spatial factor, with IQ being more strongly related than SES to differences in mean factor scores. Again, there was little difference in homogeneity of factor scores among the subpopulations.

On the perceptual speed factor there were differences, related to differences in SES, only at the low IQ level. The mean factor scores of the two high IQ subpopulations were identical and their variances virtually the same and relatively small. Both low IQ groups were widely dispersed with the effect most marked for the high SES group. In the discussion of the intercorrelation of factors it was noted that the perceptual speed factor was relatively more independent of the other factors for the two high SES subpopulations. In the high IQ-low SES subpopulation, for example, it was more highly correlated with the verbal mechanics factor, a factor with which IQ is a more highly related variable than SES. On the perceptual speed factor, more than any of the other three, care needs to be taken in applying the same substantive interpretation for all subpopulations. Although both high IQ groups have similar dispersions and the same mean factor scores, this factor does not appear to be the same. There appears to be a higher verbal component for the low SES group. There is a similar confounding of this factor in the two low IQ groups.

References

Box, G. E. A general distribution theory for a class of likelihood criteria. Biometrika, 1949, 36, 317-346.

Cochran, W. G. The X^2 test of goodness of fit. Annals of Mathematical Statistics, 1952, 23, 315-345.

Flanagan, J. C. et al. Design for a study of American youth. Boston: Houghton-Mifflin, 1962.

Gulliksen, H. and Tukey, J. W. Reliability for the Law of Comparative Judgment. Psychometrika, 1958, 23, 95-110.

Johnson, M. Factorial invariance of African educational abilities and aptitudes. Research Bulletin 69-3. Princeton, N. J.: Educational Testing Service, 1969.

Jöreskog, K. G. Some contributions to maximum likelihood factor analysis. Psychometrika, 1967, 32, 443-482. (a)

Jöreskog, K. G. UMLFA--A computer program for unrestricted maximum likelihood factor analysis. Research Memorandum 66-20 (Revised Ed.). Princeton, N. J.: Educational Testing Service, 1967. (b)

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

Jöreskog, K. G. Simultaneous factor analysis in several populations. Psychometrika, 1971, 36, in press.

Jöreskog, K. G., Gruvaeus, G. T., and van Thillo, M. ACOVS--A general computer program for analysis of covariance structures. Research Bulletin 70-15. Princeton, N. J.: Educational Testing Service, 1970.

Jöreskog, K. G. and Lawley, D. N. New methods in maximum likelihood factor analysis. British Journal of Mathematical and Statistical Psychology, 1968, 21, 85-96.

Meredith, W. Notes on factorial invariance. Psychometrika, 1964, 29, 177-186. (a)

Meredith, W. Rotation to achieve factorial invariance. Psychometrika, 1964, 29, 187-206. (b)

Rock, D. A. and Freeberg, N. E. Factorial invariance of biographical factors. Multivariate Behavioral Research, 1969, 4, 195-210.

Tucker, L. R and Lewis, C. A reliability coefficient for maximum likelihood factor analysis. Topics in Factor Analysis II, Technical Report, Office of Naval Research, in press.

van Thillo, M. and Jöreskog, K. G. SIFASP--A general computer program for simultaneous factor analysis in several populations. Research Bulletin 70-62. Princeton, N. J.: Educational Testing Service, 1970.

Table 1

Rescaled Dispersion Matrices

Group 1 Low IQ-Low SES Above Diagonal
Group 2 Low IQ-High SES Below Diagonal

	1	2	3	4	5	6	7	8	9	10	11	12
1	1.095 1.380	.743	.493	.228	.298	.381	.294	.128	.060	.150	.137	.099
2	1.038	.904 1.410	.583	.284	.351	.414	.306	.144	.063	.178	.135	.110
3	.788	.974	.776 1.324	.216	.283	.318	.214	.095	.424	.148	.128	.113
4	.261	.286	.300	.904 .934	.348	.386	.039	.003	-.023	.139	.213	.041
5	.282	.359	.280	.385	.969 1.127	.508	.197	.113	.138	.254	.179	.157
6	.368	.356	.363	.295	.389	1.303 1.465	.205	.131	.091	.258	.228	.175
7	.320	.420	.252	-.036	.162	.186	.972 1.270	.396	.359	.138	.102	.276
8	.080	.149	.081	.010	.112	.072	.302	1.128 1.255	.377	.224	.146	.338
9	.160	.236	.124	-.048	.103	.114	.498	.385	.926 1.162	.210	.097	.291
10	-.013	-.029	.002	.004	.020	.007	.050	.271	.063	1.614 2.021	.427	.483
11	.074	.091	.147	.255	.172	.131	.048	.357	.082	.384	1.450 1.737	.470
12	.033	.018	.034	-.041	-.025	-.062	.123	.379	.215	.696	.425	1.336 1.787

Table 1 (continued)

Group 3 High IQ-Low SES Above Diagonal
Group 4 High IQ-High SES Below Diagonal

	1	2	3	4	5	6	7	8	9	10	11	12
1	.947 .882	.761	.641	.273	.316	.296	.192	.099	.003	.079	.070	.029
2	.659	.968 .982	.818	.359	.408	.359	.232	.134	.010	.108	.108	.054
3	.562	.750	1.129 1.091	.334	.382	.331	.174	.117	.003	.106	.135	.080
4	.270	.303	.348	1.048 1.095	.409	.289	-.023	.015	-.062	.086	.151	.027
5	.283	.390	.368	.461	.986 .997	.378	.092	.103	.034	.130	.130	.070
6	.193	.238	.247	.242	.303	.656 .667	.127	.081	.035	.085	.123	.061
7	.252	.424	.281	-.026	.160	.088	.895 .972	.411	.440	.069	.095	.204
8	.077	.139	.134	.035	.087	.049	.316	.900 .835	.383	.075	.115	.222
9	.157	.304	.192	-.021	.196	.079	.490	.321	.967 1.030	.071	.061	.191
10	.013	.018	.042	.060	.042	.011	.015	.078	.030	.284 .307	.143	.169
11	.042	.041	.090	.177	.093	.040	-.024	.117	.014	.123	.524 .486	.201
12	-.001	.014	.045	.028	.025	.002	.084	.170	.108	.143	.153	.586 .564

Table 2

Goodness of Fit of Various Unrestricted Factor Models to Data

(21 Variables)

No. of Common Factors	χ^2	df	Prob.	$\hat{\rho}$
0	57915	210	0.00	--
1	21597	189	0.00	.588
2	13079	169	0.00	.722
3	6846	150	0.00	.838
4	3548	132	0.00	.906
5	2118	115	0.00	.937
6	1145	99	0.00	.962
7	627	84	0.00	.977
8	403	70	0.00	.983

Table 3

Goodness of Fit of Various Unrestricted Factor Models to Data

(12 Variables)

No. of Common Factors	χ^2	df	Prob.	$\hat{\rho}$
0	37018	66	0.00	--
1	11753	54	0.00	.613
2	6070	43	0.00	.750
3	2272	33	0.00	.879
4	202	24	0.00	.987

Table 4

Varimax Rotated Solution for Pooled Correlation Matrix

Tests	Common Factors I	II	III	IV	Unique Variance
Vocabulary	.728	.216	.108	.032	.410
Information I	.916	.251	.189	.017	.062
Information II	.687	.271	.094	.079	.440
Spelling	.193	.560	-.113	.125	.620
Punctuation	.190	.671	.145	.074	.487
English Usage	.206	.493	.089	.089	.698
Mechanical Reasoning	.250	.042	.654	.030	.507
Visualization I	.041	.026	.500	.249	.685
Visualization II	.046	.049	.650	.083	.566
Table Reading	.036	.086	.080	.504	.730
Clerical Checking	.040	.164	.008	.515	.707
Object Inspection	.007	-.023	.230	.615	.568

Table 5a

Restricted Correlated Solution for Pooled Correlation Matrix

Tests	Common Factors I	II	III	IV	Unique Variance
Vocabulary	.652	.220	0	0	.425
Information I	.874	.231	0	0	.038
Information II	.585	.290	0	0	.453
Spelling	0*	.579	0	0	.664
Punctuation	0	.682	0	0	.536
English Usage	0	.571	0	0	.674
Mechanical Reasoning	.216	0	.602	0	.540
Visualization I	0	0	.499	.195	.676
Visualization II	0	0	.666	0	.557
Table Reading	0	0	0	.547	.701
Clerical Checking	0	.143	0	.473	.726
Object Inspection	0	0	.173	.577	.599

*All zero elements fixed by hypothesis.

Table 5b

Intercorrelation of Factors

	I	II	III	IV
I	1.000			
II	.357	1.000		
III	.252	.159	1.000	
IV	.016	.213	.189	1.000

Table 6a

Simultaneous Solution for Four Populations

Tests	Common Factors I	II	III	IV	Unique Variances Pop. 1	Pop. 2	Pop. 3	Pop. 4
Vocabulary	.642	.231	0	0	.476	.569	.349	.350
Information I	.865	.253	0	0	.022	.078	.000	.056
Information II	.576	.300	0	0	.387	.598	.437	.476
Spelling	0	.575	0	0	.608	.602	.719	.723
Punctuation	0	.683	0	0	.497	.679	.526	.530
English Usage	0	.536	0	0	.863	1.149	.362	.464
Mechanical Reasoning	.233	0	.600	0	.549	.790	.423	.458
Visualization I	0	0	.477	.212	.771	.882	.567	.575
Visualization II	0	0	.676	0	.517	.640	.487	.548
Table Reading	0	0	0	.512	1.199	1.488	.151	.187
Clerical Checking	0	.125	0	.481	1.067	1.320	.361	.342
Object Inspection	0	0	.148	.588	.748	1.090	.338	.371

Table 6b

Estimated Factor Dispersions

		Low SES I	II	III	IV	High SES I	II	III	IV
Low IQ	I	.904				1.502			
	II	.316	1.045			.339	1.018		
	III	.111	.188	.934		.293	.108	1.068	
	IV	.122	.401	.415	1.415	-.036	.068	.208	1.857
High IQ	I	.949				.961			
	II	.443	1.004			.329	.954		
	III	.045	.051	1.127		.379	.163	1.017	
	IV	.098	.260	.203	.516	-.047	.085	.039	.487

Table 7

Estimated Factor Means

	Population	I General Knowledge	II Verbal Mechanics	III Spatial	IV Speed of Perception
1	Low IQ - Low SES	-1.78	-1.51	-.91	-.57
2	Low IQ - High SES	.58	-.31	-.75	-.40
3	High IQ - Low SES	1.00	.90	.58	.52
4	High IQ - High SES	1.27	1.27	.91	.52

Table 8

Estimated Factor Correlations

		Low SES				High SES			
		I	II	III	IV	I	II	III	IV
Low IQ	I	1.000				1.000			
	II	.334	1.000			.222	1.000		
	III	.131	.193	1.000		.183	.099	1.000	
	IV	.095	.271	.314	1.000	-.013	.036	.105	1.000
High IQ	I	1.000				1.000			
	II	.465	1.000			.359	1.000		
	III	.042	.045	1.000		.388	.168	1.000	
	IV	.200	.502	.349	1.000	-.100	.183	.079	1.000