Search for a Common Factor Model to Describe a Cross-Lagged Correlation Difference

Lloyd G. Humphreys and Charles K. Parsons
University of Illinois, Urbana-Champaign

This study describes the development of a common factor model for a cross-lagged difference involving a measure of aural comprehension and an intellectual composite. The direction of the difference was that the Listening Test predicted the composite more accurately than the composite predicted the Listening Test. A complex model which allowed seemingly identical common factors appearing at different ages to be highly, but imperfectly, correlated fit the 32 × 32 table of intercorrelations of 16 variables measured at each of two time periods. The model also described quite accurately the multiple correlations on which the cross-lagged difference was based. Aural comprehension at grade 5 was equidistant from the vectors of the two general factors at grades 5 and 11; but aural comprehension at grade 11, while close to the general factor at grade 11, was outside the space created by the two general factor vectors. Individual differences in aural comprehension anticipated by several years changes in rank on the general factor as defined by standard tests.

It is axiomatic that dependable differences between the patterns of correlations shown by two variables will result in differences in their factor structures. Some recent experience in analyzing developmental data, however, indicates that it is not always easy to go from differences in the one to differences in the other. An account of the endeavor to do so is of interest both in terms of the methodologies employed and the substantive findings which were the end result. The results demonstrate the need to be cautious in applying the common factor model to longitudinal developmental data.

The difference in correlations that created the interest in finding a factorial description was between two cross-lagged correlations: A test of aural comprehension at grade 5 predicted an intellectual composite of 15 other measures at grades 7, 9, and 11 more highly than the same composite predicted the measure of aural comprehension in the same grade combinations (Atkin, Bray, Davison, Herzberger, Humphreys & Selzer, 1977b). The differences in cross-lagged correlations were highly consistent in four groups defined by sex and race (white and black). The differences were also consistent in direction in 23 out of the 24 possible grade and group combinations. Also, considering the relatively high stability of individual differences in intelligence during the age span represented in these data, the size of the cross-lagged differences tended to be rather large. The majority of the 24 possible differences were .10 or higher; the 10 smallest differences all involved the two-year lag. A spurious basis for the cross-lagged differences in terms of changes in the reliabilities and specificities of the measures from one grade to another was ruled out by appropriate

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 2, No. 2 Spring 1978 pp. 257-267

corrections. Thus the differences can be considered psychologically "real."

The common factor model is defined as one which is concerned with communalities rather than variances. It is also concerned with factors that are invariant as a function of the composition of the battery of tests analyzed. In order to obtain a potentially invariant structure, the principal components extracted from the correlation matrix with estimates of communalities in the principal diagonal typically are rotated. Factor analysts also expect rotated factor loadings to be more meaningful psychologically than unrotated ones. The search for a common factor model for a cross-lagged correlation difference, therefore, is an attempt to find meaningful factors that will describe or explain the difference and perhaps furnish insight into possible causes for that difference. A satisfactory common factor model for these data must clearly differentiate measures of aural comprehension at different grades from each other in terms of the loadings of those measures on the common factors.

The basic data were obtained as part of a longitudinal study of intellectual growth by the Educational Testing Service. The 16 different measures used are listed in Table 1. The *N*s in the four groups were as follows: black males —172; black females—215; white males—668; and white females—762. Each person in the four samples had scores available on all measures at each of the four time periods. Because the constraint of a constant *N* was imposed in computing all correlations, the particular students on whom the analyses were based constitute a selected subset of the total nationwide sample.

Factor Structures at Grades 5 and 11

Method

The same data used in the cross-lagged analyses were also used by Atkin et al. (1977a) in an investigation of factor differentiation with age and education. Four separate analyses were reported, one for each grade, for each of the four groups. Procedures included the following:

1. Estimation of communalities by means of squared multiple correlations;
2. Decisions concerning the number of common factors by inspection of the roots and by the parallel analysis criterion (Humphreys & Montanelli, 1975; Montanelli & Humphreys, 1976);
3. Rotation to oblique structure by the Binormamin criterion (Kaiser & Dickman, 1977);

Table 1
Identification of the Measures

1.	STEP[1] Mathematics	9.	TGI[3] Industrial Arts
2.	STEP Physical Science	10.	TGI Home Arts
3.	STEP Social Science	11.	TGI Physical Science, Mathematics
4.	STEP Reading	12.	TGI Biological Science
5.	STEP Listening	13.	TGI Music, Art
6.	STEP Writing	14.	TGI History, Literature
7.	SCAT[2] Verbal	15.	TGI Recreation, Entertainment
8.	SCAT Quantitative	16.	TGI Government, Public Affairs

[1] Sequential Tests of Educational Progress
[2] School and College Ability Tests
[3] Tests of General Information

Table 2
Orthogonal Factor Loadings of Listening at Grades 5 and 11

	"g"	Comprehension		Information			h^2
Grade 5	69	28		04			55
		Verbal	Quantitative	Fem.	Mas.	Soc. Sci.	
Grade 11	73	26	-03	-09	05	09	62
Grade 5 (Dwyer extension)	70	23	-11	-04	14	09	58

4. Extraction of a second order factor; and
5. Transformation of all factors into an orthogonal hierarchical order (Schmid & Leiman, 1957).

The rationale for rotations to a hierarchical structure in a sense dates back to Spearman's discussion of a general factor. The full hierarchy has been discussed in considerable detail by Vernon (1950) and Humphreys (1962). It is easy to defend a model of intelligence that includes both a general factor and group factors. The importance of a general factor is evident from the intercorrelations of cognitive measures in the full range of human talent.

Results

Factor loadings for Listening at the extreme grade levels were extracted from these analyses and are presented in the first two rows of Table 2. Only the sample of white males, who will be the subjects of the present series of factor analytic studies, are presented. Data from blacks were not used after the first cycle because those data included more sampling error "noise." Results for white females, on the other hand, were almost completely parallel to those for white males; therefore, presentation of both would be redundant. The only discernible difference between the sexes was that the factors were defined somewhat more clearly in the male sample.

Before discussing the significance of the factor data for the present problem, a brief description of the factors is in order. Only two factors were justified by the criteria used at the grade 5 level. Rotation to a hierarchical order resulted in three factors: (1) a general factor of the type described by Spearman; (2) a broad comprehension factor defined by the two SCAT tests and six STEP tests; and (3) a general information factor defined by the eight tests of highly specific information. In grade 11 five factors were congruent with the criteria used. Broad comprehension was split into verbal and quantitative factors, while the general information factor was subdivided into "feminine" (aesthetic), "masculine" (mechanical and science), and social science information. The five first-order factors defined a single second-order factor which became the general factor after the use of the transformation.

It is clear that the test of primary interest—Listening—was not represented in the factors added as age and education increase. Listening at grade 11 continued to load on both the general factor and the group factor of verbal comprehension which represented a narrowing of the grade 5 comprehension factor. There did not seem to be any basis in these loadings for the highly significant cross-lagged difference.

The third row of Table 2 shows the results of extending the grade 5 measures onto the grade 11 factor structure by means of the multiple

Table 3
Oblique Factor Loadings of Listening in a Matrix Combining Grades 5 and 11

	Comprehension	Information	Verbal	Quantitative	Fem.	Mas.	Soc. Sci.	h^2
Listening at 5	54	-03	24	-19	-12	20	18	63
Listening at 11	00	-07	66	02	-16	08	25	62

regression method described by Dwyer (1937). To explain the cross-lagged difference, one possible outcome would have been a higher general factor loading in the grade 11 analysis for grade 5 Listening than for grade 11 Listening. A second possible outcome would have been higher loadings of grade 5 Listening on the differentiated grade 11 factors. Neither occurred. The sum of the squared loadings was a little higher for the extended factor loadings than for the three factors obtained in the analysis of the grade 5 intercorrelations. However, this h^2 was no larger than the one obtained by arbitrarily extracting five principal components from those same intercorrelations. There is still no basis in this table for the cross-lagged difference.

Note that a negative outcome here does not constitute disconfirming evidence concerning the difference in correlations. That difference is real, as evidenced by its consistency and by traditional sampling error estimates. Negative outcomes of these or any other factor descriptions cast doubt on the validity of the factor methodology employed for these developmental data.

Analysis of Grades 5 and 11 Combined

Method

Although an analysis of a 64 × 64 matrix (16 measures in two or more time periods combined) was readily possible, interpretations of factors are uncertain because a simplex process underlies the intercorrelations over the four time periods for each of the 16 measures. (See Humphreys, 1960, for a discussion of this problem.) A four-variable simplex automatically determines two factors, even though there is no change in the nature of the measures that define the factors. The analysis of a 32 × 32 matrix covering only two time periods does not encounter quite the same difficulty; nevertheless, the single test-retest correlation for each of the 16 measures should contain both common and specific variance. On the basis of both experience and theory, it was expected that the presence of specific variance components of appreciable size, particularly in the tests of specific information, would complicate an analysis of a 32 × 32 matrix.

A computer program developed by Tucker and Dunn[1] avoids the above difficulty. The program belongs to the MINRES (Harman, 1967) family and is called a Weighted MINRES solution. Particular correlations can be given any weight, including zero, in the factor solution. This does not mean that the correlation in question is set equal to zero. Rather, it is given no weight in determining the unrotated factor loadings. Thus, as applied to these data, the 16 test-retest correlations were given zero weight and the number of factors set at seven: two for grade 5 and five for grade 11. Rotated loadings for Listening appear in Table 3. The standard hierar-

[1]The computer program, by L. Tucker and T. Dunn, 1976, is on file in Computing Services, University of Illinois, Urbana-Champaign.

Table 4
Comparison of Residuals from Weighted and Unweighted MINRES Factor Extraction in a 32 x 32 Matrix

	Principal Diagonal		Secondary Diagonal	
	Weighted	Unweighted	Weighted	Unweighted
Mean	.306	.304	.025	.018
Median	.307	.300	.022	.014
S.D.	.032	.031	.028	.013
Range	.162 to .516	.164 to .524	-.058 to .073	-.007 to .048

chical solution was abandoned because there were small, but systematic, discrepancies to the fit furnished by a single factor in the second order.

Results

The results were again negative. There was no basis in the factor descriptions for Listening at the two grade levels for the development of a theory concerning the cross-lagged difference. It is perhaps particularly surprising that the communalities were so nearly equal. Since Listening at grade 5 was more broadly correlated with grade 11 tests than the reverse, a difference in communalities might be expected.

There was one incidental finding of general interest. Although test-retest correlations were all given zero weight in computing the unrotated loadings, the seven common factors accounted for most of the variance in those correlations. When the analysis was rerun with the unweighted MINRES program, factors were almost identical. A comparison of the residuals from the two solutions appears in Table 4.

The mean and median residuals in both the principal diagonal and the secondary diagonal (test-retest correlations) of the matrix differed only slightly. The difference was sufficiently small that it could be explained by the difference in the number of correlations determining the factors, i.e., 16 fewer correlations were used in the weighted MINRES analysis. There was a larger difference between the two analyses in the variability of the residuals in the secondary diagonal which was caused almost entirely by Mathematics and Quantitative Aptitude. The quantitative factor was defined by only those two measures; they defined it primarily at the grade 11 level. The omission of the two test-retest correlations allowed a degree of indeterminancy in the factor loadings. It is also possible that the tests loaded on the quantitative factor are not as closely parallel at the two grade levels as are the other tests.

The psychological significance of the near identity of the factors in the two sets of computations is that over a period of six years of relatively rapid intellectual development, specific factor scores in these tests are almost completely unstable. For purposes of long term prediction, the specific factors, although included in true score variance at any one time period, function in much the same way as error variance. Under certain circumstances it would be legitimate, for theoretical purposes, to correct correlations for the attenuation introduced by the combination of error and nonerror specifics (unique variance).

Analyses of Measures at Mixed Grade Levels

It was noted in the discussion of the 32 × 32 variable analysis that the factor structure in the

Table 5
Comparison of General Factor Loadings and Communalities of Factoring, Verbal, and Quantitative Aptitude in Eight Analyses

	Grade 5 Analyses							
	All Grade 5		Listening 11		Verbal 11		Quantitative 11	
	"g"	h^2	"g"	h^2	"g"	h^2	"g"	h^2
Listening	73	55	60	37	75	63	72	56
Verbal Aptitude	87	81	87	79	79	68	85	78
Quantitative Aptitude	75	64	74	63	73	66	62	47
	Grade 11 Analyses							
	All Grade 11		Listening 5		Verbal 5		Quantitative 5	
	"g"	h^2	"g"	h^2	"g"	h^2	"g"	h^2
Listening	73	62	71	57	73	61	73	62
Verbal Aptitude	89	84	89	87	75	62	89	84
Quantitative Aptitude	72	71	70	72	71	72	63	51

second order was quite complex. There was evidence for communality associated with a particular time of testing. There was also evidence for communality between time periods resulting from the factor differentiation. This complexity suggested the next analyses which involved the substitution, one by one, of grade 5 measures for their grade 11 counterparts and vice-versa in a series of analyses at each of the two grade levels. While it seemed unnecessary to do this systematically for each of the 16 measures, some control for its use with Listening was required. Both Verbal and Quantitative Aptitude were treated in this way for control purposes. The resulting general factor loadings and communalities are presented in Table 5 along with loadings from standard analyses based on all 16 variables at a particular grade level.

Note that each entry in the table is from a separate factor analysis. Four were conducted at grade 5 and four at grade 11. The first entry (or column) at each level is the one based on factoring all 16 measures at that level. The second entry is from matrices in which Listening at 11 was substituted for Listening at 5 or in which Listening at 5 was substituted for Listening at 11. The third and fourth entries are from matrices in which Verbal and Quantitative Aptitude, respectively, were switched in the same fashion. If the loading of a variable does not change across a row in the table, a test at a given grade can be substituted for the same test at the other grade without change in factor composition. If the loading drops, the substitute is less adequate at the other grade level.

Inspection of these loadings indicates that there was a clear difference between the loadings of Listening at grades 5 and 11 which was not true, with equal clarity, for either Verbal or Quantitative Aptitude. Listening at grade 5 had a loading on the general factor at grade 11 which was hardly distinguishable from Listening at grade 11. The reverse was clearly not true. Listening at grade 11 had a substantially reduced loading on the general factor at grade 5. Loadings for Verbal and Quantitative Aptitude showed a drop in both directions. Similar statements can be made about the communalities.

Is there a factorial explanation for the series of negative findings followed by the positive results in the present analyses? The explanation is inherent in the simplex process: The general factor at grade 5 is not identical with the general factor at grade 11. Lack of identity can be of two types: in the content of the measures or in the instability of individual differences in the factor scores. As long as psychological tests have to be adapted to the ability levels of the examinees, a choice between these alternative hypotheses remains logically in doubt. The measurement of stature is a different matter. There is no change in the scale of measurement, but the intercorrelations of stature at several time periods formed a simplex matrix. Four time periods will define two common factors, six will define three, and so on. Each of the factors clearly measures stature, but the persons measured change their rank order in the group as they mature.

The next problem was to determine whether or not the explanation phrased in terms of lack of factor identity from one time period to another could be demonstrated in an analysis of a 32 × 32 matrix; Jöreskog's Restricted Maximum Likelihood program (1970) was the only factor analytic method that would allow the test of a model as complex as believed necessary.

Restricted Maximum Likelihood Analysis

If it is hypothesized that the general factor is not identical at two different time periods, it also seems highly probable that group factors would not be identical. This assumption further complicates the factor model, which must allow for all factors that can be defined by the 16 tests at the grade levels selected for analysis. To develop a model for 16 tests at four time periods would have been extremely complex; it was decided, therefore, to use only the two extreme grades as had been done in the analysis of the 32 × 32 matrix previously reported. It had also been found in the previous analyses that a poorly defined Quantitative factor could be identified at grade 5. These decisions alone resulted in the need to define six factors in the 32 × 32 matrix, i.e., two general, two verbal comprehension, and two quantitative. The addition of the general information factor at grade 5 and the three differentiated information factors at grade 11 resulted in a total of 10 factors hypothesized to describe the intercorrelations of 32 measures.

Since it had already been determined that seven factors were sufficient to describe the intercorrelations, the model proposed is not parsimonious. If parsimony were a primary goal, however, there would be no use for the common factor model. Extraction without rotation from a variance-covariance matrix (either standardized or unstandardized) of the r principal components that describe a high but arbitrary percentage of the total variance in the n variables would be eminently satisfactory on the basis of the criterion of parsimony. It is the psychological meaning of rotated factors determined from the common variance among the measures which is of interest. Furthermore, there is a criterion for the accuracy of the model: It must be able to reproduce a highly stable cross-lagged difference involving Listening and the remainder of the measures.

Work was begun with the Jöreskog program on the sample of white females. Because more unknowns had to be determined than there were degrees of freedom, work proceeded in an iterative fashion, fixing in subsequent iterations values that had been determined in a preceding one. Clearly, there is some degree of capitalization on chance in a procedure of this type. Thus, when a seemingly good fit was obtained for the females, the data for the males were analyzed. Now it was possible to start the process by fixing initially more values than it was possible to do in the first sample. After only five iterations with the new sample, χ^2 was reduced to 440 with 387 degrees of freedom. Relative to the size of N, the fit was highly acceptable. For example, χ^2 for seven principal components was 484.

In contrast to the usual order of presentation, the intercorrelations of the factors are presented prior to the factor loadings. This is done because

Table 6
Intercorrelations of the Factors Following
Restricted Maximum Likelihood Analysis

		1	2	3	4	5	6	7	8	9	10
"g" at 5	1		81	0	08	0	-02	0	16	-01	13
"g" at 11	2			11	0	03	0	08	0	0	0
Verbal at 5	3				87	0	10	0	21	13	23
Verbal at 11	4					12	0	13	0	0	0
Quantitative at 5	5						80	0	11	25	14
Quantitative at 11	6							25	0	0	0
Information at 5	7								46	19	32
Feminine at 11	8									0	0
Masculine at 11	9										0
Social Science at 11	10										

the factor intercorrelations represent the important aspect of the common factor model required to represent these developmental data. They are shown in Table 6. A single zero indicates a correlation that was fixed by the model. All others were allowed to vary. Thus, the orthogonal hierarchical rotations used at each grade level separately were carried over to the model. Correlations between group factors at one age and the general factor at a different age still tended to be close to zero. Correlations between the general, verbal, and quantitative factors at different ages were high, but appreciably less than unity. The multiple correlation between the three information factors at grade 11 and the general information factor at grade 5 was only moderately high. The information factors were seemingly less stable than the general and comprehension factors.

Since factors that are described and named in identical fashion are not identical at different age levels, a factor analysis that imposes identity is bound to lead to psychological conclusions that are inaccurate. Note, however, that these data do not enable us to distinguish between two alternative explanations for the lack of identity. The tests defining the factors may be functionally different at the two ages, or the persons in the sample may have factor scores that are imperfectly correlated over time. Given the facts that every effort was made to make these tests comparable from grade school through high school and that development is rapid during this time period, the alternative, phrased in terms of change in the students being examined, appears to be the more plausible of the two explanations.

Factor loadings for Listening and five other measures from the total of 16 are shown in Table 7. The five added to Listening for comparative purposes were selected to represent all of the grade 11 factors. The findings for these measures did not differ in any way from those for the ones omitted. Before interpreting these loadings, it is necessary to recall the nature of oblique factor loadings which are not represented by perpendicular projections. Oblique projections can be deceiving, particularly when factors are highly correlated. For example, Listening at grade 5 was conspicuous in the set in terms of its moderate loadings on both general factors; yet the vector of Listening at grade 5 was about the same length as the vector of Listening at 11 in the space defined by the two general factors. Listening at grade 5 was equidistant from the two general factors; it was an equally good measure of both. Listening at 11 was a good measure of

Table 7
Factor Loadings of Six Representative Variables Following Restricted Maximum Likelihood Analysis

	1	2	3	4	5	6	7	8	9	10
Math 5	81	-01	00	06	35	03	00	-05	-03	-01
Math 11	-01	73	-02	03	-02	44	00	-05	-06	-06
Listening 5	34	37	15	16	03	06	-09	18	11	17
Listening 11	-10	77	10	27	16	-04	05	05	-05	10
Verbal 5	85	-01	19	-01	04	03	17	00	05	-03
Verbal 11	01	87	03	19	01	03	00	06	13	13
Ind. Arts 5	71	01	-05	02	01	-05	43	13	01	-10
Ind. Arts 11	-01	50	-02	-07	02	-03	01	40	07	-05
Home Arts 5	63	-01	-05	04	05	-04	45	01	03	-01
Home Arts 11	-04	69	-06	00	-01	-04	00	04	20	14
Government 5	66	01	07	-01	-05	03	40	-09	-06	10
Government 11	01	66	01	03	-05	-02	03	-02	01	34

the general factor at 11, but it had a negative loading on the grade 5 general factor. It was more distant from the general factor at 5, as a matter of fact, than any of the other grade 11 measures.

It is this general factor loading difference that makes the principal contribution to the cross-lagged difference and is, in all probability, the most dependable source of that difference in the sampling sense. Small contributions to the cross-lagged difference in this sample were also made by the loadings on the three information factors at grade 11 since, in every case, the loading of Listening at grade 5 was larger than the loading of its counterpart at grade 11. Differences here, however, were smaller than those on the general factor.

Even though the intercorrelations were reproduced from the model within reasonable error tolerances, small, consistent errors in zero-order correlations might add up to appreciable errors in the multiple correlations used in the cross-lagged analysis. To check on this possibility, all of the multiple correlations for the same six variables appearing in Table 7 were computed from the *R* matrix determined by the model. Comparisons between predicted and observed multiple correlations are presented in Table 8.

Columns 1 and 4 contain the synchronous correlations, and columns 2 and 3 contain the cross-lagged values. By subjective standards the fit is quite good. While there seem to be cross-lagged differences for certain of the tests other than Listening, such as Mathematics and Industrial Arts information, these largely disappeared when corrections were made for changes in the uniqueness of the measures between the two time periods. The difference for Listening was affected only slightly by the same correction. In addition, whether corrected or uncorrected, no other measure, including the two mentioned, showed the consistency from group to group and from one time period to another as did Listening. It can be concluded that the model does describe the basic phenomena.

Conclusions

The model is admittedly complex, but this appears to be necessary. In fact, it is so complex that it was decided not to attempt bringing in to the common factor model the data from grades

Table 8
Comparison of Predicted and Observed
Synchronous and Cross-Lagged Multiple Correlations

	1		2		3		4	
	Predicted	Observed	Predicted	Observed	Predicted	Observed	Predicted	Observed
Mathematics	835	848	726	728	643	648	792	807
Listening	755	749	736	730	633	612	787	770
Verbal Aptitude	874	886	771	764	836	808	908	894
Industrial Arts	800	807	674	654	471	491	542	559
Home Arts	749	759	608	605	600	591	657	673
Government	760	774	660	656	634	623	705	709

1 = Test at Grade 5, Composite at Grade 5
2 = Test at Grade 5, Composite at Grade 11
3 = Test at Grade 11, Composite at Grade 5
4 = Test at Grade 11, Composite at Grade 11

7 and 9. The common factor approach may generally be of limited utility in dealing with developmental observations. Nevertheless, an important finding is represented in the demonstration here that individual differences in aural comprehension at grade 5 anticipate, within the limits of reliability and specificity of the test, individual differences in general intelligence at grade 11. The difficulties encountered in developing the model that led to this conclusion support a methodological conclusion as well. Use of traditional factor analytic methodology for developmental data will produce factors and factor loadings that are misleading psychologically. The factor model must allow for change.

Most clinicians are convinced that an individual test of intelligence is more valid than a group test. It is a reasonable inference that this conviction is based upon the difference in mode of comprehension rather than the difference between individual and group administration. The Listening Test in the STEP series is not the most reliable of the set. It is also a good deal less reliable than the Verbal Aptitude section of SCAT. Listening could be made more reliable, and thus a more valid measure of general intelligence, by lengthening the test and by controlling more carefully its mode of administration. It might also be made more valid by the introduction of additional types of aural comprehension tasks. When the scores on a test anticipate by several years individual differences in general intelligence as measured by printed tests, the test itself clearly deserves the attention of those concerned with psychometric research and development.

References

Atkin, R., Bray, R., Davison, M., Herzberger, S., Humphreys, L., & Selzer, U. Ability factor differentiation, grades 5 through 11. *Applied Psychological Measurement*, 1977, *1*, 65–76. (a)

Atkin, R., Bray, R., Davison, M., Herzberger, S., Humphreys, L., & Selzer, U. Cross-lagged panel analysis of sixteen cognitive measures at four grade levels. *Journal for Research in Child Development*, 1977, *48*, 944–952. (b)

Dwyer, P. S. The determination of the factor loadings of a given test from the known factor loadings of other tests. *Psychometrika*, 1937, *2*, 173–178.

Harman, H. H. *Modern factor analysis* (2nd ed. rev.) Chicago: The University of Chicago Press, 1967.

Humphreys, L. G. Investigations of the simplex. *Psychometrika*, 1960, *25*, 313–323.

Humphreys, L. G. The organization of human abilities. *American Psychologist*, 1962, *17*, 475–483.

Humphreys, L. G., & Montanelli, R. G., Jr. An investigation of the parallel analysis criterion for determining the number of common factors. *Multivariate Behavioral Research*, 1975, *10*, 193–205.

Jöreskog, K. G. A general method for analysis of covariance structures, *Biometrika*, 1970, *57*, 239–511.

Kaiser, H. R., & Dickman, K. W. Some properties of binormamin. *Psychometrika*, 1977, *42*, 307–309.

Montanelli, R. G., Jr., & Humphreys, L. G. Latent roots of random data correlation matrices with squared multiple correlations on the diagonal; A monte carlo study. *Psychometrika*, 1976, *41*, 341–348.

Schmid, J., & Leiman, J. The development of hierarchical factor solutions. *Psychometrika*, 1957, *22*, 53–61.

Vernon, P. E. *The Structure of human abilities.* New York: Wiley, 1950.

Acknowledgements

This research was supported by a grant from the National Institute of Mental Health, MH 23612-04, Studies of Intellectual Development and Organization. We thank Dr. Thomas L. Hilton and Educational Testing Service for providing data from the Study of Academic Growth and Prediction.

Author's Address

Lloyd G. Humphreys, 425 Psychology Building, University of Illinois, Champaign, IL 61820.