Aging and Structural Invariance in Intelligence[1]

Christopher K. Hertzog[2] and K. Warner Schaie
Andrus Gerontology Center and
Department of Psychology
University of Southern California

A paper presented at the 87th Annual Convention of the American Psychological Association, New York City, September 1979.

INTRODUCTION

A major issue in the study of adult intellectual development has been whether the factor structure of intelligence remains qualitatively invariant with advancing age. Age-related invariance in factor structures would provide evidence that the observed or manifest variables, the psychometric tests, are measuring the same latent intellectual factors at different ages; indeed, Baltes and Nesselroade (1970, 1973) argued that a demonstration of structural invariance is necessary before quantitative age changes in mean performance levels can be interpreted unambiguously.

Garrett (1946), Anastasi (1970) and others have interpreted factor analytic results as indicating evidence for dedifferentiation of the intellectual factor structure from young adulthood to senescence (see Reinert, 1970). Dedifferentiation in its most extreme form implies a qualitative change in the underlying factor space, even to the point of suggesting a fewer number of common factors in older subject's data due to collapsing hyperplanes, eventually leading to a single intelligence factor (G?) accounting for all observed variables' common variance. Evidence for the dedifferentiation hypothesis exists, although the studies supporting the hypothesis have usually found only a trend towards a collapsing factor space; however, the studies favoring dedifferentiation can be countered by studies supporting an invariance hypothesis (e.g., Riegel & Reigel, 1962). Much of the contradictory evidence may be due to differing factor analytic techniques and criteria for invariance (Cunningham, 1978; Reinert, 1970).

A theoretical paper by Meredith (1964) bears directly upon the issue of appropriate criteria for assessing group differences in factor structure. Specifically, Meredith (1964) used Lawley's selection theorem to show that, if a factor analysis model holds for a given population, then selection of subgroups from that population should still yield an invariant factor pattern matrix of raw score regressions of manifest variables on factors. However, the covariance matrices of manifest variables, unique components, and factors would not generally be equivalent across groups. Meredith's (1964) paper is important with respect to the dedifferentiation hypothesis because 1) it suggests that age differences in standardized factor loadings or in factor covariance matrices would be expected by age selection alone, and cannot be taken as evidence of qualitative age differences, and 2) only variation in the raw score factor pattern matrix constitutes evidence of qualitative age differences in factor structure (Mulaik, 1972).

Recent advances in restricted maximum likelihood factor analysis techniques by Jöreskog and coworkers (e.g. Jöreskog, 1969, 1971; Sörbom & Jöreskog, 1976) are directly applicable to the problem of testing hypotheses of structural invariance in multiple groups (see Bechtoldt, 1974; McGaw & Jöreskog, 1971). These methods are also extremely useful in testing hypotheses of interindividual stability in longitudinal factor analysis models (e.g., Jöreskog and Sörbom, 1977). These methods are generally preferable to other longitudinal factor analysis models (e.g., Corballis and Traub, 1970) precisely because they directly estimate raw score regressions of manifest variables on factors and the factor covariance matrix (as opposed to the factor correlation matrix), thus enabling separate tests of 1) cross-occasion invariance in factor loadings, 2) cross-occasion changes in magnitude of individual differences (reflected in factor variances), and 3) cross-occasion stability in interindividual differences (reflected in factor covariances). The factor analysis models of Jöreskog and Sörbom (1977) are also particularly suited for longitudinal analysis because they allow for nonzero covariances between unique and covariance between identical measures over occasions is likely in longitudinal data (Sörbom, 1975). Omission of these autocorrelated residuals would perturb the estimated factor loading and factor covariances.

The present study used the longitudinal models of Jöreskog and Sörbom (1977) to test hypotheses of between-group and cross-occasion invariance of factor pattern, factor covariance, and unique covariance matrices in two longitudinal sequences from Schaie's Seattle study.

METHOD

Factor Analysis Model

All models were estimated from the following general factor analysis model (Sörbom & Jöreskog, 1976): Given a px1 vector of observed variables x_g in group g, with mean vector μ_g and covariance matrix Σ_g, then a factor analysis model with a k x 1 vector of common factors f_g and a p x 1 vector of unique factors z_g is

$$x_g = \nu_g + \Lambda_g f_g = z_g, \qquad (1)$$

where ν_g is a p x 1 vector of grand means and Λ_g is a p x k matrix of raw score regressions of x_g on f_g. Then the factor means, a k x 1 vector Θ_g relates to μ_g by

$$\mu_g = \nu_g + \Lambda_g \Theta_g \qquad (2)$$

The structural model is

$$\Sigma_g = \Lambda_g \Phi_g \Lambda_g + \Psi_g \qquad (3)$$

where Φ is the covariance matrix of f_g and Ψ is thecovariance of matrix z_g.

Parameters in $\Theta_g, \Lambda_g, \Phi_g$ and Ψ_g are fixed, free, or constrained to equal other unknown parameters (f_g and z_g are not estimated). Provided that the investigator has specified a unique model by constraining of fixing parameters, the COFAMM program (Sörbom and Jöreskog, 1976) will estimate all unknown parameters and their standard errors of estimate, while also providing a χ^2 goodness of fit test for the fit of the estimated Σ_g, $\hat{\Sigma}_g$, to S_g, the sample covariance matrix. Fitting of $\hat{\Sigma}_g$ to S_g is accomplished by minimizing a fitting function with respect to all unknown parameters (which is equivalent to maximizing a log likelihood function). The χ^2 goodness of fit statistic may be used to compare improvement in fit from a parent model to a less restricted counterpart (i.e., one with additional free parameters), by evaluating the difference in corresponding χ^2 statistics (Jöreskog 1971, 1974).[3] COFAMM also provides the first derivatives of all fixed and constrained parameters, wich may be used to identify parameters which fit the data poorly (Sörbom, 1975).

Data from two 14 year longitudinal samples were used in the present report. The first sample consisted of 162 men and women from seven 7-year birth cohorts, tested on three occasions (1956, 1963, 1970). The second sample, consisted of 250 men and women from 7-year birth cohorts, also tested on three occasions (1963, 1970, 1977). Longitudinal factor analyses were performed on each sample, ignoring the cohort and sex classifications. In addition, the two samples were pooled to form a larger sample for simultaneous multiple group analysis. Three groups were formed by combining across samples for matched age intervals (see Table 1). The data matrix does not permit unambiguous interpretation with respect to age, cohort, and time effects (although the effects are not completely confounded), but this was deemed necessary in order to increase sample sizes enough to justify large sample assumptions in the factor analysis.

<u>Variables</u>

The variables were the five subtests of Thurstone's 1949 version of the Primary Mental Abilities intelligence battery: Verbal Meaning (V), a test of recognition vocabulary; space (S), a test of figural discrimination under two dimensional rotation; Reasoning, (R) a test of inductive reasoning by letter series completion; Number (N), a test of speeded two column addition; and word fluency (W), a test of speeded retrieval from semantic memory. There were 15 dependent measures in the analyses, representing the five PMA subtests at each of the three measurement occasions.

Models

Two basic models were estimated: 1) an occasion-specific model of the type studied by Jöreskog and Sörbom (1977), and 2) a test-specific model. The occasion-specific model fit a general (G) factor at each occasion in $\underset{\sim}{\Lambda}$, and left $\underset{\sim}{\Phi}$ unconstrained. The selection of a G factor representation for the 5 subtests was indicated by an exploratory factor analysis of first occasion data for 2200 subjects. The test-specific model fit 5 factors to the data, one for each PMA subtest. Although the occasion-specific model is the preferred model, in that it explicitly models time-dependent changes among corresponding parameters in $\underset{\sim}{\Lambda}$, $\underset{\sim}{\Phi}$ and $\underset{\sim}{\Psi}$, the test-specific model corresponds to the one likely to be obtained in an exploration factor analysis, and the test-specific $\underset{\sim}{\Phi}$ provides information about covariances among primary abilities. The $\underset{\sim}{\Lambda}$ matrices for the two basic models are shown in Figure 1.

RESULTS[4]

The procedure consisted of testing occasion-specific and test-specific models in Sample 1, replicating the best-fitting models in Sample 2, and testing multiple group models in the pooled sample.

Occasion-Specific Models

The initial occasion-specific model in Sample 1, H_1 specified a G factor with 5 free λ elements at each occasion, $\underset{\sim}{\Phi}$ scaled to a correlation matrix by fixing G variances to unity, and a diagonal $\underset{\sim}{\Psi}$ matrix of unique variances. The model fit poorly (See Table 2), even though all parameters estimated were large relative to their standard errors. However, offdiagonal $\underset{\sim}{\Phi}$ elements exceeded unity. Perturbations due to autocorrelated unique components were probably contributing to the poor fit, and H_2 allowed autocorrelated residuals for all occasions, with a marked improvement in fit (indeed, absolute χ^2 was no longer significant).

Two other models tested cross-occasion invariance in $\underset{\sim}{\Lambda}$. H_3 was identical to H_2 but constrained corresponding λ elements to equivalence across occasions. The factors were still standardized separately by fixing diagonal λ elements to unity; H_3 therefore required the standardized regressions of variables on G (G factor loadings) to be equivalent across occasions. An alternative model, H_4, defined the metric of G by fixing factor loadings of R on G to unity and left $\underset{\sim}{\Phi}$ unconstrained; H_4 therefore constrains the unstandardized G factor loadings to cross-occasion equivalence without allowing cross-occasion differences in G factor variances to affect the hypothesis of cross-occasion invariance in $\underset{\sim}{\Lambda}$. The results from these models indicate H_4 to fit better than H_3 (Table 2), but H_4 did not fit better than H_2. Model H_4 modeling cross-occasion invariance in unstandardized factor loadings, but allowing cross-occasion variability in $\underset{\sim}{\Phi}$ and $\underset{\sim}{\Psi}$ was therefore accepted as the most reasonable model. Table 3 gives the parameters, standard errors, and scaled solution values for the accepted model. G was defined primarily by R and V, as can be seen from the residual variances and the scaled factor loadings ($\underset{\sim}{\Lambda}*$)[5]; although all variables load appreciably on G, the unique variances for S, N, and W are relatively large. There was little change in G variance (diagonal ϕ elements) between the first two occasions, but there was a substantial increase in G variance between the second and third occasions. The off-diagonal elements in $\underset{\sim}{\Phi}^*$ show G to be highly stable across occasions, implying stability in the distribution of individuals about the G factor means.

The model specification for H_4 was well replicated for Sample 2, for although the absolute χ^2 was significant, the fitting function value was smaller than that achieved in Sample 1. 95% confidence intervals about the parameter estimates for both samples overlapped in all cases.

Test-Specific Models

A test-specific model was computed for Sample 1 by specifying 1) 3 non-zero elements in each of the k=5 columns of $\underset{\sim}{\Lambda}$, one for each replicated subtest on its corresponding test-specific factor; 2) standardized oblique factors in $\underset{\sim}{\Phi}$, with variances fixed to unity and freely estimated off-diagonal factor correlations; and 3) $\underset{\sim}{\Psi}$ to be restricted to a diagonal matrix of 15 uncorrelated unique variances. Results are given in Table 4. The model fit relatively well, with large factor loadings and small unique variances, but the χ^2 test was still significant. The factors were highly correlated, particularly V and R.

The salient first derivatives were associated with residual covariances in $\underset{\sim}{\Psi}$. A sequential relaxation of fixed zero covariances (as recommended by Sörbom, 1975) failed to improve fit to the level achieved by the occasion-specific model unless non-significant parameters were allowed. Thus the occasion-specific model achieves a level of fit with theoretically meaningful parameters which cannot be approximated by the test-specific model.

In spite of this fact, the properties of the test-specific model were of theoretical interest; hence the model was replicated for Sample 2. Again excellent replication was obtained, with 95% confidence intervals overlapping for all parameters. An interesting tendency in $\underset{\sim}{\Phi}$ was that the correlation between V and R was slightly lower, and the correlation between R and S was higher for Sample 2. However, the parameters were so similar to the previous model they are not reported here.

Multiple Group Analyses

Analyses treating the two sequential samples as single groups had proved useful in model building, and had indicated a high degree of longitudinal stability in the data. Longitudinal stability was reflected in the occasion-specific models as high cross-occasion correlation in G and in cross-occasion invariance in G factor loadings; longitudinal stability was reflected in the test-specific model as large and consistent loadings of variables on test-specific factors.

Stability in the single group analyses does not imply stability in all subsamples, however. In particular, collapsing over the entire age/cobort range may serve to obscure structural metamorphosis in the oldest age range, where the number of subjects is smallest. The simultaneous sultiple groups analyses were designed to address this issue.

An initial test of the equality of $\underset{\sim}{\Sigma}_g$ revealed significant group differences (Box's M = 402.77; F = 1.59 $(240,\infty)$, $p <.0001$). Thus group differences in some factor analytic parameters seemed probable.

Occasion-Specific Models. A sequence of occasion-specific models were tested. The first model, H_1, required 1) both between-group and cross-occasion invariance in G factor loadings; 2) an unrestricted $\underset{\sim}{\Phi}$ matrix, with parameters constrained to between-group invariance; and 3) a non-diagonal $\underset{\sim}{\Psi}$ matrix of the type estimated for H_2 in Sample 1, constrained to between group invariance. H_1 is denoted:

$$H_1: \quad \underset{\sim}{\Lambda}_t = \quad \underset{\sim}{\Lambda}_g = \quad \underset{\sim}{\Phi}_g = \quad \underset{\sim}{\Psi}_g = \quad .$$

A sequence of models relaxed some of these restrictions. H_2 relaxed the constraint of cross-occasion invariance in G factor loadings:

$$H_2: \Lambda_g =, \Phi_g =, \Psi_g =; \Lambda_t \neq .$$

The third model, H_3, relaxed between-group constraints on Φ_g:

$$H_3: \Lambda_g =, \Psi_g =, \Lambda_t \neq, \Phi_g \neq$$

H_4 relaxed the constraints on Ψ:

$$H_4: \Lambda_g =; \Lambda_t \neq, \Phi_g \neq, \Psi_g \neq .$$

H_5 relaxed all equivalence constraints:

$$H_5: \Lambda_t \neq, \Lambda_g \neq, \Phi_g \neq, \Psi_g \neq .$$

Finally, H_6 reimposed constraints on Λ:

$$H_6: \Lambda_t =, \Lambda_g =; \Phi_g \neq, \Psi_g \neq .$$

Table 5 gives the goodness of fit statistics for models $H_1 = H_6$. In general, results showed improvement in fit when Φ_g and Ψ_g differed across groups. but no improvement when constraints on Λ were relaxed. The most parsimonious model therefore appeared to be H_6, which allowed group differences in Φ_g and Ψ_g but fitted a single regression matrix of PMA subtests on G for all occasions and groups. All free parameters were significantly non-zero (i.e., greater than twice their standard error). Tables 6 and 7 give the parameter estimates and scaled solution values.

The following important points can be seen in these results:

1) Parameters in Λ followed the previously observed pattern: G factor loadings were largest for R and V_1 and smaller for N_1 W_1 and S (in decreasing

order of magnitude).

2) Differences in Φ_g appeared to reflect group differences in G variance (diagonal elements). G variance is smallest for the youngest group (Group 1) and largest for the oldest group (Group 3).

3) The pattern of cross-occasion changes in G variance differed between groups; there was some decrease in G variance in Group 1, relative stability in Group 2 (albeit with some increase between the second and third occasions), and a substantial increase in variability for Group 3 between the second and third occasions.

4) The within-group stability in individual differences, as reflected in covariance elements of Φ_g, was uniformly high; when Φ was separately rescaled to a correlation matrix for all groups the correlations exceeded .88 (Table 8).

5) The scaled unique variances in Ψ_g, shown in Table 7, show a general tendency towards to decrease from Group 1 to Group 3, indicating higher communalities for Group 3.

A potential confound exists with regard to Group 3; namely, that it was formed by pooling over a wider age/cohort range. This wider range may have produced the larger variances in Φ_3. This hypothesis was tested by forming a new Group 3, using only the 67 subjects in the two oldest cohorts. The G variances were still reliably larger, but the cross-occasion increase in G variance was virtually eliminated. Nevertheless, the greater variance in G for Group 3 could not be due merely to its wider age/cohort range.

Test-specific Models. A similar sequence of test-specific models was tested. The sequence was initiated by testing the hypothesis that all groups had the same number of common factors, using separate unrestricted maximum likelihood factor analysis on each group. Table 9 gives the goodness of fit tests. As expected 5 factors were clearly indicated for Groups 1 and 2; 5 factors were also indicated for Group 3, although the data were more equivocal.

The first restricted test-specific model required between-group equivalence in all matrices:

$$H_1: \Lambda_g = , \Phi_g = , \Psi_g = .$$

Subsequent models relaxed some or all of these constraints, as for the occasion-specific models. These models and their associated fits are shown in Table 10. Model H_4, allowing group differences in Φ_g and Ψ_g, but forcing between-groups invariance in Λ_g, was the accepted model. The main parameters of interest here were the elements of Φ_g, showing group differences in PMA factor variances and covariances. Table 11 gives the scaled solution values, and Table 12 shows Φ_g separately scaled as correlation matrices. The correlations among PMA factors are relatively similar in Groups 1 and 2 (except for an increased correlation between V and R). In Group 3, however, the factor correlations are uniformly higher and all significantly non-zero. Given the high correlation between V and R, a sixth model, H_6 in Table 10, fitted 4 factors in Group 3, requiring V and R subtests for form a single factor. The model did not fit as well as H_5; and the hypothesis of only 4 common factors as specified in H_6 was rejected.

<u>Occasion-Specific Model with Factor Means</u>. One of the chief advantages of the occasion-specific model is that, the factor means, when estimated, reflect cross-occasion and between group differences in performance level expressed in terms of the factors rather than observed variables. An occasion-specific estimating factor means in Θ_g was estimated, using Model H_6 from the occasion-specific sequence (i.e., H_6: $\Lambda_t = , \Lambda_g = ; \Phi_g \neq , \Psi_g \neq$).

A problem arises in COFAMM applications to longitudinal factor means, namely that the location parameters (grand means) in ν are free to vary over occasions, which implies that some of the longitudinal means differences will be absorbed in $\underset{\sim}{\nu}$ and not represented in $\underset{\sim}{\Theta}_g$. Fortunately, the observed means for Group 2 showed little longitudinal variation; hence fixing $\underset{\sim}{\Theta}_2$ to a 3 x 1 vector of zeroes (one group must have fixed zero elements in $\underset{\sim}{\Theta}$ to identify $\underset{\sim}{\Theta}_g$; see Sörbom 1974) resulted in substantially invariant $\underset{\sim}{\nu}$ across occasions. The parameter estimates and standard errors are given in Table 13, and the means are graphed in Figure 2. The overall fit of the model was not as good as obtained for H_6 above, indicating that adding $\underset{\sim}{\Theta}_g$ had reduced the goodness of fit. The differences in G means were substantial, conforming to expectations from Schaie's previous analyses (Schaie, 1979). Group 1 performed at a higher level than Groups 2 and 3, and showed increment from ages 30-44. Group 2 was forced to stability (but the residuals indicated a good fit of the means). Group 3, on the other hand, performed at much lower levels than Groups 1 and 2 and showed substantial decline in performance levels, particularly between the last two occasions. The means for the redefined older group, Group 3* in Figure 2, further substantiated the decline in performance level in old age. Note also the substantial level differences between groups at the endpoints (e.g., Occasion 1 for Group 3, Occasion 3 for Group 2) where mean ages were virtually idenfical. This pattern is suggestive of generational differences in performance level.

Discussion

The results of this study are not consistent with the hypothesis of substantive qualitative structural change in intelligence over the adult life span; to the contrary, the structure of intelligence in the models tested here shows signs of relative stability. The major finding was invariance in the raw score regressions of PMA variables on factors in $\underset{\sim}{\Lambda}$, both across occasions and between groups in the occasion-specific models, and between groups in the test-specific models. The finding of between-group invariance in both G factor loadings in the occasion-specific model and PMA factor loadings in the test-specific model is consistent with the hypothesis that the age/cohort groups are selected from a single population in which a common factor analysis model holds (Meredith, 1964), and implies that the structural differences in Φ and Ψ reflect group selection and not qualitative differences in factor structure. There is, moreover, evidence of cross-occasion invariance in G factor loadings in the occasion-specific models, indicating that structural invariance holds at the intraindividual level as well. Taken together, these results suggest that the PMA subtests are indeed measuring the same latent factors within and between individuals of different ages, at least in the somewhat select longitudinal samples studied here.

Under the assumption that cohort differences in factor structure are minimal, there was some indication of a modest age-dedifferentiation process, as reflected in higher communalities for Group 3 in the occasion-specific model and in a higher factor intercorrelations for Group 3 in the test-specific model, which is consistent with other findings (e.g., Cunningham, Note 1; Cunningham and Birren, in press). Similar findings led Cunningham and Birren (in press) to suggest the influence of an age-related slowing in cognitive speed. The present results would seem to indicate that such slowing, if it is indeed the effect underlying age selection, had not resulted in major qualitative differences in the mapping of PMA variables onto their associated factors (which is parsimonious, since the PMA tests were directly intended to tap both speed and power components; Thurstone and Thurstone, 1941; 1949). It is possible, of course, that the trend toward dedifferentiation found in this study could, if continued, ultimately lead to qualitative changes in the factor structure, but this had apparently not occurred in the present data.

There is, in general, evidence for longitudinal stability in the structural models tested in this study. Longitudinal stability was reflected in high covariances among G factors across occasions in the occasion-specific model, implying stability in the ordering of individual differences over a 14 year period. Admittedly, such stability may be specific to the type of long term longitudinal sample studied here, which is known to be influenced by a high rate of experimental mortality (e.g., Schaie, Labouvie, and Barrett, 1973) and may not generalize to the population at large. Nevertheless, the magnitude of the cross-occasion covariances in $\underset{\sim}{\Phi}_g$ was impressively high.

While there is evidence of structural stability, the occasion-specific model with factor means indicates that not all groups show stability in level of performance: the younger group showed mean increment, the middle-aged group showed mean stability, and the older group showed mean decline; this overall pattern is of course consistent with previous reports on level differences by Schaie and coworkers (see Schaie, 1979). These results again support the notion that mean intraindividual decrement in intelligence as measured by the PMA begins no earlier than the fifth decade of life; stability in performance levels is seen prior to this point.

While these interpretations are not new, of course, the critical point is that they have been based on data from models which satisfy the suggestions of Baltes and Nesselroade (1970, 1973) to demonstrate structural invariance while investigating age changes in performance level.

An interesting variant in the stability/decline issue may be indicated in the occasion-specific pattern of G variances in $\underset{\sim}{\Phi}_g$. In the single sample analysis, large increases in G variance were found between the second and third occasions, but this increase was found to be primarily specific to Group 3 in the multiple group analyses. Furthermore, this effect was eliminated when Group 3 was redefined by eliminating two cohort groups with mean ages ranging from 53-67, which altered the mean ages of Group 3 from 58-72 to 63-77 over the

14 year longitudinal interval. One possible interpretation is that there is a general transition from a stability pattern to a decremental pattern in intra-individual levels of performance from roughly age 55 to age 70, and which was reflected in increasing G variance over occasions when Group 3 was formed by pooling middle-aged subjects with old subjects. This hypothesis is admittedly speculative and in need of a more rigorous test.

Finally, the results of this study point to the power and utility of Jöreskog and Sörbom's restricted maximum likelihood factor analytic techniques for longitudinal factor analysis. In the present application, where G variances differed over occasions, the estimation of standardized factors would have led to the conclusion that the factor loadings varied over occasions. This study suggests that it is only the standardized (or scaled) factor loadings which vary, since the raw score loadings could be taken as invariant, and this pattern of results has a substantially different interpretation with regard to structural invariance than would habe been obtained using standardized factors. Thus COFAMM (and LISREL) seem particularly well suited for the analysis of longitudinal data.

Reference Note

1. Cunningham, W.R. Factorial variation in intellectual ability tests of intermediate speed and difficulty. Paper presented at the 31st Annual Scientific Meeting of the Gerontological Society; Dallas, Texas, 1978.

REFERENCES

nastasi, A. On the formation of psychological traits. American Psychologist, 1970, 25, 899-910.

Baltes, P.B. and Nesselroade, J.R. Multivariate longitudinal and cross-sectional sequences for analyzing ontogenetic and generational change. Developmental Psychology, 1970, 2, 163-168.

Baltes, P.B. and Nesselroade, J.R. The developmental analysis of individual differences on multiple measures. In J.R. Nesselroade and H.W. Reese (Eds.), Life Span Developmental Psychology: Methodological Issues. New York, Academic Press, 1973, 219-252.

Bechtoldt, H.P. A Confirmatory analysis of the factor stability hypothesis. Psychometrika, 1974, 39, 319-326.

Corballis, M.C and Traub R.E. Longitudinal factor analysis. Psychometrika, 1970, 35, 79-93.

Cunningham, W.R. Principles for identifying structural differences: Some methodological issues related to comparative factor analysis. Journal of Gerontology, 1978, 33, 82-86.

Cunningham, W.R. and Birren, J.E. Age changes in the factor structure of intellectual abilities in adulthood and old age. Educational and Psychological Measurement, (in press).

Garrett, H.E. A developmental theory of intelligence. American Psychologist, 1946, 1, 372-378.

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, 409-426.

Jöreskog, K.G. Analyzing psychological data by analysis of covariance structures. In D.H. Krantz, R.C. Atkinson, R.D. Luce, and P. Suppes (Eds.) Contemporary Developments in Mathematical Psychology, (Vol. 2). San Francisco, W.H. Freeman, 1974, pp. 1-56.

Jöreskog, K.G. and Sörbom, D. Statistical models and methods for analyses of longitudinal data. In D.S. Aigner and A.S. Goldberger (Eds.) Latent Variables in Socio-economic Models. Amsterdam, North Holland, 1977, pp. 285-325.

McGaw, B. and Jöreskog, K.G. Factorial invariance of ability measures in groups differing in intelligence and socioeconomic status. British Journal of Mathematical and Statistical Psychology, 1971, 24, 154-168.

Meredith, W. Notes on factorial invariance. Psychometrika, 1964, 29, 177-185.

Mulaik, S.A. Foundations of Factor Analysis. New York, McGraw-Hill, 1972.

Reinert, G. Comparative factor analytic studies of intelligence throughout the human life span. In L.R. Goulet and P.B. Baltes (Eds.), Life Span Developmental Psychology: Research and Theory. New York, Academic Press, 1970, pp. 115-145.

Riegel R.M. and Riegel, K.F. A comparison and reinterpretation of factor structures of the W-B, the WAIS, and the HAWIE on aged persons. Journal of Consulting Psychology, 1962, 26, 31-37.

Schaie, K.W. The Primary Mental Abilities in Adulthood: An exploration in the development of psychometric intelligence. In P.B. Baltes and O.G. Brim, Jr. (Eds), Life Span Development and Behavior. Vo. 2. New York, Academic Press, 1979, pp. 67-115.

Schaie K.W., Labouvie, G.V. and Barrett, T.J. Selective attrition effects in a 14-year study of adult intelligence. Journal of Gerontology, 1973, 28, 328-334.

Sörbom, D. A general method for studying differences in factor means and factor structure between groups. British Journal of Mathematical and Statistical Psychology, 1974, 27, 229-239.

Sörbom, D. Detection of correlated errors in longitudinal data. British Journal of Mathematical and Statistical Psychology, 1975, 28, 138-151.

Sörbom D. and Jöreskog, K.G. COFAMM: Confirmatory Factor Analysis with Model Modification, Chicago, Noational Educational Resources, 1976.

Thurstone, L.L. and Thurnstone, T.G. Factorial studies of intelligence. Psychometric Monographs, No. 2. Chicago, University of Chicago Press, 1941.

Thurstone, L. L. and Thurstone, T.G. Examiner Manual for the SRA Primary Mental Abilities Test. Chicago, Science Research Associates, 1949.

FOOTNOTES

1. This paper reports part of the analyses from Dr. Hertzog's doctoral dissertation. We thank William Meredith for his helpful suggestions on the models testing cross-occasion invariance in factor pattern matrices.

2. Present location: Department of Psychology, University of Washington.

3. The χ^2 goodness of fit test is a product of the number of subjects and the fitting function (F) at minimum. Hence absolute χ^2 may be significant in large samples even with relatively good fit, and should not be taken as the sole criterion for accepting a model (Jöreskog, 1971).

4. In the interest of brevity, several models tested are not presented in this paper, and the parameters of other models alluded to in this paper are sometimes not reported. Interested individuals are urged to contact the first author for additional details.

5. The scaled solution is computed by weighting Λ_g, ϕ_g, and Θ_g by the square root of the pooled factor variance for each element. The advantage is that scaled factor loadings are of the same magnitude as standardized factor loadings (and reduce to standardized factor loadings with a single group) while the scaled Φ_g is on average a correlation matrix (i.e. $E(\text{diag}\Phi_g) = 1$). Explicit scaling formulae are given in Jöreskog (1971) and Sörbom and Jöreskog, (1976).

Table 1

Reparameterized Sequential Sample for
Multiple Group Analysis

	Sample	Cohort (mean birth year)	Age (means)	N
Group 1			30, 37, 44	109
	1	1931	25, 32, 39	21
	1	1924	32, 39, 46	26
	2	1938	25, 32, 39	22
	2	1931	32, 39, 46	40
Group 2			42, 49, 56	160
	1	1917	39, 46, 53	27
	1	1910	46, 53, 60	32
	2	1924	39, 46, 53	51
	2	1917	46, 53, 60	50
Group 3			58, 65, 72	143
	1	1903	53, 60, 67	28
	1	1896	60, 67, 74	15
	1	1889	67, 74, 81	13
	2	1910	53, 60, 67	48
	2	1903	60, 67, 74	18
	2	1896	67, 74, 81	21

Table 2

Summary of Goodness of Fit Tests:

Occasion-Specific Models

Model	X^2	df	P	F	Comparisons	ΔX^2	df	p
H_1	988.89	87	.000	3.070	-	-	-	-
H_2	84.13	72	.155	.261	$H_1 - H_2$	904.76	15	$<.001$
H_3	99.69	82	.090	.310	$H_3 - H_4$	6.41	2	$<.05$
H_4	93.28	80	.147	.290	$H_2 - H_4$	9.16	8	$>.05$

Table 3
Accepted Occasion-Specific
Model (H_4)

Λ[a]		Λ_t[b]	Λ^*[c]		
			G_1	G_2	G_3
	V	.935 (.061)[d]	.801	.801	.871
	S	.645 (.069)	.553	.553	.601
	R	1[e] (-)	.857	.856	.932
	N	.610 (.071)	.523	.523	.569
	W	.596 (.072)	.511	.510	.555

Φ		G_1	G_2	G_3
	G_1	.734 (.104)		
	G_2	.697 (.098)	.733 (.104)	
	G_3	.735 (.102)	.777 (.106)	.868 (.116)

Φ[f]		G_1	G_2	G_3
	G_1	1		
	G_2	.950	1	
	G_3	.921	.973	1

[a]Fixed zero elements of matrix omitted.

[b]Occasion-specific factor loadings (invariant over occasions).

[c]Scaled factor loadings on occasion-specific factors (G_1, G_2, G_3).

[d]Standard errors in parentheses.

[e]Fixed parameter.

[f]Scaled factor correlation matrix.

[g]Unique variances for subtests (number denotes occasion).

[h]Covariances of unique components of identical subtests across occasions (numbers denote occasions).

Table 3 (Cont'd.)

Ψ^a	Var 1[g]	Var 2[g]	Var 3[g]
V	.319 (.051)	.382 (.056)	.271 (.046)
S	.762 (.089)	.635 (.075)	.634 (.075)
R	.256 (.052)	.259 (.051)	.128 (.041)
N	.708 (.083)	.775 (.090)	.690 (.080)
W	.808 (.094)	.786 (.091)	.693 (.080)
	Cov 12[h]	Cov 13[h]	Cov 23[h]
V	.141 (.043)	.141 (.040)	.172 (.043)
S	.448 (.069)	.406 (.067)	.431 (.064)
R	.105 (.043)	.062 (.038)	.082 (.039)
N	.593 (.078)	.605 (.076)	.590 (.077)
W	.562 (.080)	.498 (.074)	.518 (.074)

[a]Fixed zero elements of matrix omitted.

[b]Occasion-specific factor loadings (invariant over occasions).

[c]Scaled factor loadings on occasion-specific factors (G_1, G_2, G_3).

[d]Standard errors in parentheses.

[e]Fixed parameter.

[f]Scaled factor correlation matrix.

[g]Unique variances for subtests (number denotes occasion).

[h]Covariances of unique components of identical subtests across occasions (numbers denote occasions).

Table 4

Test-Specific Model

	V	S	R	N	W
$\tilde{\Lambda}$[a]					
Occasions					
1	.862 (.064)[b]	.783 (.068)	.870 (.063)	.928 (.060)	.834 (.066)
2	.884 (.063)	.919 (.063)	.909 (.061)	.892 (.062)	.885 (.064)
3	.920 (.061)	.869 (.065)	.952 (.059)	.951 (.059)	.866 (.065)
$\tilde{\Phi}$[c]					
V	1				
S	.564 (.061)	1			
R	.869 (.026)	.671 (.050)	1		
N	.502 (.064)	.309 (.077)	.534 (.061)	1	
W	.553 (.063)	.238 (.083)	.511 (.065)	.423 (.071)	1
$\tilde{\Psi}$[d]					
Occasions					
1	.256 (.036)	.386 (.051)	.242 (.032)	.138 (.024)	.304 (.046)
2	.218 (.033)	.156 (.038)	.175 (.026)	.204 (.029)	.217 (.042)
3	.154 (.028)	.245 (.041)	.094 (.020)	.096 (.023)	.250 (.043)

[a]Fixed zero elements omitted; each subtest loads only on a subtest-specific factor (e.g., V_1, V_2, V_3 on V).

[b]Standard errors in parentheses.

[c]Diagonal elements fixed to unity.

[d]Covariances (off-diagonal elements) were fixed to sero and omitted from table.

Table 5

Summary of Goodness of Fit Statistics for Multiple Groups Analyses: Occasion-Specific Models

Model	χ^2	df	p	F	Comp[a]	$\Delta\chi^{2}$[b]	df	p
H_1: Λ_t = ,Λ_g = ,Φ_g = ,Ψ_g =	514.60	320	.000	.629				
H_2: Λ_g = ,Φ_g = ,Ψ_g = ;$\Lambda_t \neq$	506.65	312	.000	.619	H_1-H_2	7.95	8	> .05
H_3: Λ_g = ,Ψ_g = ;$\Lambda_t \neq$,$\Phi_g \neq$	453.53	300	.000	.554	H_2-H_3	53.12	12	< .001
H_4: Λ_g = ;$\Lambda_t \neq$,$\Phi_g \neq$,$\Psi_g \neq$	284.24	240	.026	.347	H_3-H_4	80.33	12	< .001
H_5: $\Lambda_t \neq$,$\Lambda_g \neq$,$\Phi_g \neq$,$\Psi_g \neq$	257.85	216	.027	.315	H_4-H_5	26.39	24	> .05
H_6: Λ_t = ,Λ_g = ,$\Phi_g \neq$,$\Psi_g \neq$	291.49	248	.030	.356	H_5-H_6	33.64	32	> .05
					H_6-H_4	7.25	8	> .05

[a]Column indicates models compared by testing differences in χ^2 for statistical significance.

[b]Difference in χ^2 between two models under comparison.

Table 6

Results from Accepted Multiple Groups
Occasion-Specific Model (H_6)

Λ[a]	G
V	.988 (.058)[b]
S	.573 (.053)
R	1[c] (-)
N	.781 (.057)
W	.631 (.056)

Φ_g[d]		G_1	G_2	G_3
Φ_1				
	G_1	.468 (.089)		
	G_2	.371 (.075)	.374 (.076)	
	G_3	.372 (.074)	.333 (.068)	.341 (.070)
Φ_2				
	G_1	.523 (.084)		
	G_2	.503 (.079)	.518 (.082)	
	G_3	.525 (.082)	.522 (.081)	.572 (.088)
Φ_3				
	G_1	.735 (.112)		
	G_2	.700 (.107)	.746 (.116)	
	G	.737 (.113)	.804 (.120)	.941 (.137)

[a]Occasion-specific factor loadings (invariant over groups and occasions); fixed zero elements omitted.

[b]Standard errors in parentheses.

[c]Fixed parameters.

[d]Factor covariance matrix (subscripts denote group).

[e]Residual covariance matrix (subscripts denote group).

[f]Residual variances (unique variances); number denotes occasion.

[g]Covariance of residual (unique) components of identical subtests; numbers denote occasions.

Table 6 (Cont'd.)

Ψ_g[e]

Ψ_1

	Var 1[f]	Var 2[f]	Var 3[f]
V	.411 (.078)	.333 (.067)	.360 (.067)
S	.953 (.135)	1.078 (.151)	1.099 (.154)
R	.385 (.077)	.500 (.088)	.415 (.076)
N	.691 (.105)	.717 (.107)	.673 (.100)
W	.721 (.105)	.920 (.131)	.976 (.138)
	Cov 12[g]	Cov 13[g]	Cov 23[g]
V	.287 (.064)	.270 (.064)	.257 (.060)
S	.814 (.129)	.783 (.128)	.797 (.134)
R	.212 (.067)	.197 (.063)	.264 (.069)
N	.548 (.095)	.575 (.094)	.554 (.094)
W	.466 (.096)	.453 (.097)	.712 (.119)

[a]Occasion-specific factor loadings (invariant over groups and occasions); fixed zero elements omitted.

[b]Standard errors in parentheses.

[c]Fixed parameters.

[d]Factor covariance matrix (subscripts denote group).

[e]Residual covariance matrix (subscripts denote group).

[f]Residual variances (unique variances); number denotes occasion.

[g]Covariance of residual (unique) components of identical subtests; numbers denote occasions.

Table 6 (Cont'd)

Ψ_2	Var 1[f]	Var 2[f]	Var 3[f]
V	.367 (060)	.412 (.064)	.335 (.057)
S	.929 (.108)	.784 (.092)	.804 (.094)
R	.481 (.072)	.389 (.062)	.388 (.063)
N	.716 (.088)	.792 (.091)	.648 (.081)
W	.824 (.098)	.688 (.082)	.604 (.073)
	Cov 12[g]	Cov 13[g]	Cov 23[g]
V	.269 (.055)	.211 (.051)	.227 (.052)
S	.582 (.086)	.447 (.081)	.502 (.078)
R	.270 (.058)	.270 (.058)	.245 (.054)
N	.609 (.083)	.565 (.078)	.572 (.079)
W	.512 (.077)	.464 (.072)	.463 (.068)

[a]Occasion-specific factor loadings (invariant over groups and occasions); fixed zero elements omitted.

[b]Standard errors in parenthese.

[c]Fixed parameters.

[d]Factor covariance matrix (subscripts denote group).

[e]Residual covariance matrix (subscripts denote group).

[f]Residual variances (unique variances); number denotes occasion.

[g]Covariance of residual (unique) components of identical subtests; numbers denote occasions.

Table 6 (Cont'd)

Ψ_3	Var 1[f]	Var 2[f]	Var 3[f]
V	.441 (.071)	.565 (.083)	.477 (.075)
S	.703 (.088)	.652 (.081)	.558 (.071)
R	.331 (.061)	.379 (.065)	.260 (.055)
N	.587 (.078)	.625 (.082)	.514 (.070)
W	.805 (.101)	.792 (.099)	.681 (.086)
	Cov 12[g]	Cov 13[g]	Cov 23[g]
V	.164 (.060)	.161 (.056)	.245 (.064)
S	.293 (.065)	.331 (.063)	.297 (.060)
R	.157 (.051)	.116 (.046)	.167 (.050)
N	.464 (.071)	.391 (.064)	.405 (.066)
W	.559 (.086)	.509 (.080)	.508 (.080)

[a]Occasion-specific factor loadings (invariant over groups and occasions); fixed zero elements omitted.

[b]Standard errors in parentheses.

[c]Fixed parameter.

[d]Factor covariance matrix (subscripts denote group).

[e]Residual covariance matrix (subscripts denote group).

[f]Residual variances (unique variances); number denotes occasion.

[g]Covariance of residual (unique) components of identical subtests; numbers denote occasions.

Table 7

Scaled Solution for Multiple Group Occasion-Specific Model

			G_1	G_2	G_3
Λ^{*a}					
		V	.754	.739	.790
		S	.437	.428	.458
		R	.763	.748	.799
		N	.596	.584	.625
		W	.481	.472	.504
			G_1	G_2	G_3
Φ^*_g	Φ^{*b}_1				
		G_1	.804	.670	.534
		G_2	.650	.558	
		G_3	.609		
	Φ^{*b}_2				
		G_1	.899	.926	.894
		G_2	.881	.874	
		G_3	.861		
	Φ^{*b}_3				
		G_1	1.262	1.334	1.473
		G_2	1.226	1.344	
		G_3	1.209		

[a] Fixed zero elements omitted.

[b] Subscript denotes group.

[c] Matrix of unique variances (scaled separately for each group). Unique variances have been rescaled as a proportion of the estimated population variances from $\hat{\Sigma}_g$ (variances of observed variables for each group are equal to 1.0).

Table 7 (Cont'd)

Ψ^*_g[c]	V	S	R	N	W
Ψ^*_1[b]					
Occasion					
1	.474	.862	.451	.707	.795
2	.477	.898	.572	.759	.861
3	.519	.908	.549	.763	.878
Ψ^*_2[b]					
Occasion					
1	.418	.845	.479	.692	.798
2	.449	.822	.429	.701	.770
3	.375	.810	.404	.650	.727
Ψ^*_3[b]					
Occasion					
1	.381	.745	.311	.567	.734
2	.437	.727	.337	.579	.727
3	.342	.644	.216	.472	.645

[a]Fixed zero elements omitted.

[b]Subscript denotes group.

[c]Matrix of unique variances (scaled separately for each group) Unique variances have been rescaled as a proportion of the estimated population variances from $\hat{\Sigma}_g$ (variances of the observed variables for each group are equal to 1.0).

Table 8

Occasion-Specific Model:
Rescaled Correlation Matrices

		Occasion		
$\underset{\sim}{\Phi}_1^{**}$[a]		G_1	G_2	G_3
	G_1	1		
	G_2	.886	1	
	G_3	.929	.933	1
$\underset{\sim}{\Phi}_2^{**}$				
	G_1	1		
	G_2	.966	1	
	G_3	.960	.961	1
$\underset{\sim}{\Phi}_3^{**}$				
	G_1	1		
	G_2	.945	1	
	G_3	.887	.952	1

[a]Correlation matrix among G factors (subscript denotes group).

[b]Correlation matrix among unique (residual) elements (subscript denotes group). Diagonal unities and fixed zero correlations omitted.

[c]Correlation among residuals of same subtest at different occasions (subscripts denote occasions).

Table 8 (Cont'd)

Ψ_1^{**}[b]		r_{12}[c]	r_{13}[c]	r_{23}[c]
	V	.776	.831	.742
	S	.803	.898	.732
	R	.483	.618	.580
	N	.779	.929	.798
	W	.572	.632	.751
Ψ_2^{**}				
	V	.692	.731	.611
	S	.682	.655	.632
	R	.624	.787	.631
	N	.836	.883	.825
	W	.680	.751	.718
Ψ_3^{**}				
	V	.329	.490	.472
	S	.433	.724	.492
	R	.443	.493	.532
	N	.766	.802	.715
	W	.700	.796	.692

[a]Correlation matrix among G factors (subscript denotes group).

[b]Correlation matrix among unique (residual) elements (subscript denotes group). Diagonal unities and fixed zero correlations omitted.

[c]Correlation among residuals of same subtest at different occasions (subscripts denote occasions).

Table 9

χ^2 Statistics for Number of Factors

Number of Factors	χ^2	df	$\Delta\chi^2$	df	T-L
			Group 1		
1	797.35	90	-	-	.30
2	532.77	76	264.58	14	.47
3	290.15	63	242.62	13	.68
4	164.54	51	125.61	12	.80
5	74.76	40	89.78	11	.92
6	45.78	30	28.98	10	.95
7	26.38	21	19.40	9	.98
			Group 2		
1	1072.12	90	-	-	.40
2	294.29	76	377.83	14	.55
3	425.01	63	269.18	13	.68
4	199.60	51	225.51	12	.84
5	38.18	40	161.42	11	1.00
6	20.17	30	17.47	10	1.02
7	9.48	21	11.23	9	1.03
			Group 3		
1	659.56	90	-	-	.59
2	465.98	76	193.58	14	.67
3	237.82	63	228.16	13	.82
4	142.37	51	95.45	12	.88
5	80.38	40	61.99	11	.93
6	35.58	30	44.80	10	.99
7	11.88	21	23.70	9	1.03

Abbreviations: df = degrees of function;
$\Delta\chi^2$ = change in χ^2;
T-L = Tucker-Lewis reliability coefficient

Table 10

Summary of Goodness of Fit Tests:
Multiple Groups; Test-Specific Models

Model	χ^{2} [a]	df	F	Comp[b]	$\Delta\chi^2$	df	p
H_1: Λ_g = , Φ_g = , Ψ_g =	589.76	320	.721				
H_2: Λ_g = , Ψ_g = , $\Phi_g \neq$	543.67	290	.665	H_1-H_2	46.09	30	< .05
H_3: Λ_g = , Φ_g = , $\Psi_g \neq$	447.70	290	.547	H_1-H_3	142.07	30	< .001
H_4: Λ_g = ; $\Phi_g \neq$, $\Psi_g \neq$	398.85	260	.488	H_3-H_4	48.85	30	< .05
H_5: $\Lambda_g \neq$, $\Phi_g \neq$, $\Psi_g \neq$	371.07	240	.454	H_4-H_5	27.78	20	> .05
H_6: $\Lambda_g \neq$, $\Phi_g \neq$, $\Psi_g \neq$ K=4 in Group 3	433.68	244	.530	H_6-H_5	62.61	4	.001

[a] All absolute χ^2 significant beyond .001 level.

[b] Model comparisons made by computed differences in corresponding χ^2 statistics.

Table 11

Multiple Groups Test-Specific Model:
Scaled Solution

Λ^{*a}		V	S	R	N	W
Occasions						
1		.895	.812	.861	.921	.808
2		.891	.864	.889	.914	.883
3		.864	.808	.912	.916	.873
Φ_1^{*b}						
	V	.698				
	S	.111	1.345			
	R	.408	.461	.762		
	N	.324	.276	.425	.990	
	W	.468	.042	.392	.351	1.114
Φ_2^{*b}						
	V	.974				
	S	.284	.944			
	R	.719	.477	1.051		
	N	.407	.237	.445	.967	
	W	.475	.036	.443	.287	.899
Φ_3^{*b}						
	V	1.259				
	S	.595	.800			
	R	.997	.616	1.124		
	N	.764	.483	.679	1.045	
	W	.633	.263	.543	.466	1.027

[a] Scaled test-specific factor loadings.

[b] Subscript denotes group; the weighted average of Φ_g is a correlation matrix.

[c] Subscript denotes group; unique variances expressed as a proportion of estimated observed variances from Σ_g (observed variances scaled to unity).

Table 11 (Cont'd.)

Ψ_1^c	V	S	R	N	W
Occasion					
1	.250	.183	.304	.162	.374
2	.140	.195	.329	.212	.209
3	.190	.272	.219	.109	.257
Ψ_2^c					
Occasion					
1	.137	.192	.226	.137	.373
2	.165	.199	.162	.149	.219
3	.199	.402	.147	.141	.196
Ψ_3^c					
Occasion					
1	.274	.445	.199	.165	.323
2	.282	.368	.182	.153	.231
3	.271	.363	.161	.206	.241

[a]Scaled test-specific factor loadings.

[b]Subscript denotes group; the weighted average of $\underset{\sim}{\Phi}_g$ is a correlation matrix.

[c]Subscript denotes group; unique variances expressed as a proportion of estimated observed variances from $\underset{\sim}{\Sigma}_g$ (observed variances scaled to unity).

Table 12

Test-Specific Model[a]: Standardized Φ_g

Φ_1^{**}[b]	V	S	R	N	W
V	1				
S	.115	1			
R	.559	.455	1		
N	.390	.239	.489	1	
W	.531	.034	.425	.334	1
Φ_2^{**}[b]					
V	1				
S	.296	1			
R	.711	.479	1		
N	.419	.248	.441	1	
W	.508	.039	.439	.308	1
Φ_3^{**}[b]					
V	1				
S	.593	1			
R	.838	.650	1		
N	.666	.528	.627	1	
W	.557	.290	.505	.450	1

[a] Φ_g^{**} defined as Φ_g scaled to correlation matrix (Φ_g standardized separately for each group.

[b] Subscript denotes group.

Table 13

Occasion-Specific Model:
Model with Factor Means[a]

Θ_g	G_1	G_2	G_3
Θ_1[b]	.268 (.107)[c]	.452 (.102)	.501 (.102)
Θ_2	0 (-)[d]	0 (-)	0 (-)
Θ_3	-.729 (.109)	-.836 (.110)	1.152 (.117)
μ[e]	1	2	3
V	4.252	4.295	4.151
S	2.312	2.357	2.335
R	2.982	2.972	2.994
N	2.570	2.557	2.563
W	3.750	3.752	3.690
Λ[f]			
	.878 (.040)		
	.656 (.041)		
	1.000 (-)		
	.566 (.040)		
	.522 (.041)		

[a] Ψ matrices omitted.

[b] Subscript denotes group.

[c] Standard errors in parentheses.

[d] Means for group 2 fixed at 0.

[e] Matrix of grand means (invariant over groups).

[f] Occasion-specific factor loadings on G (invariant over groups and occasions).

Table 13 (Cont'd.)

Φ_1			
1	.527 (.101)		
2	.416 (.086)	.417 (.087)	
3	.420 (.083)	.375 (.078)	.386 (.080)
Φ_2			
1	.616 (.094)		
2	.592 (.088)	.609 (.092)	
3	.619 (.091)	.614 (.090)	.673 (.098)
Φ_3			
1	.811 (.119)		
2	.759 (.112)	.824 (.123)	
3	.793 (.118)	.879 (.125)	1.020 (.142)

[a] Ψ matrices omitted.

[b] Subscript denotes group.

[c] Standard errors in parentheses.

[d] Means for group 2 fixed at 0.

[e] Matrix of grand means (invariant over groups).

[f] Occasion-specific factor loadings on G (invariant over groups and occasions).

Figure 1

Factor Loadings (Λ) for Longitudinal Factor Analysis Models

A. Occasion-Specific Λ

	G_1	G_2	G_3
V_1	λ_1	0	0
S_1	λ_2	0	0
R_1	λ_3	0	0
N_1	λ_4	0	0
W_1	λ_5	0	0
V_2	0	λ_6	0
S_2	0	λ_7	0
R_2	0	λ_8	0
N_2	0	λ_9	0
W_2	0	λ_{10}	0
V_3	0	0	λ_{11}
S_3	0	0	λ_{12}
R_3	0	0	λ_{13}
N_3	0	0	λ_{14}
W_3	0	0	λ_{15}

Figure 1 continued

B. Test-Specific $\underset{\sim}{\Lambda}$

	V	S	R	N	W
V_1	λ_1	0	0	0	0
S_1	0	λ_2	0	0	0
R_1	0	0	λ_3	0	0
N_1	0	0	0	λ_4	0
W_1	0	0	0	0	λ_5
V_2	λ_6	0	0	0	0
S_2	0	λ_7	0	0	0
R_2	0	0	λ_8	0	0
N_2	0	0	0	λ_9	0
W_2	0	0	0	0	λ_{10}
V_3	λ_{11}	0	0	0	0
S_3	0	λ_{12}	0	0	0
R_3	0	0	λ_{13}	0	0
N_3	0	0	0	λ_{14}	0
W_3	0	0	0	0	λ_{15}

Figure 2

G Factor Means from $\underset{\sim}{\theta}_g$