15

ON THE CHANGE IN MENTAL ORGANIZATION WITH AGE:

AN ALTERNATIVE APPROACH

Hans-Magne Eikeland
University of Oslo

Oslo, May 1973

This report is a preliminary version issued for limited circulation. Corrections, criticisms, and suggestions for revision are solicited. The report should not be cited as a reference without the specific permission of the author.

ON THE CHANGE IN MENTAL ORGANIZATION WITH AGE: AN ALTERNATIVE APPROACH.

Hans-Magne Eikeland, University of Oslo

Introduction.

In the 'thirties and 'forties intelligence research was much concerned with studying change in the structure of mental abilities as a function of age (e.g., Asch 1936, Clark 1944, Garrett, Bryan, and Perl 1935, Garrett 1938, Reichard 1944). Garrett's (1938, 1946) wellknown age-differentiation hypothesis maintained that as age increases the influence of the socalled g-factor decreases, and more specialized abilities will come to dominate.

Many contradictory results were obtained during the years on that hypothesis. Therefore, no conclusive evidence seems ever to have been reached that settles the controversy about it. Vernon (1950) is though fairly well assured that the relatively many confirmatory results from U.S. studies can be attributed to an artificially increasing homogeneity of samples with age. This is brought about by comparing college students with high school students, and high school students with elementary school children. These groups represent populations that reflect an increasing selection with age. Hence it is reasonable to believe that the smaller g-variance in the older groups is due to their greater selectivity:

> "...the writer would conclude that there is no general tendency towards differentiation, except perhaps in early infancy;

> and that everything depends on the type of educational and vocational training. Usually where abilities are practised at school or in jobs they tend to become more specialized, though sometimes teaching is of such a character as to increase integration. Again regression or de-differentiation may occur as effects of past training wear off. It is conceiv-able that secondary schooling is more fragmentary in America than in Britain, and so apt to produce more differentiation between 12 and 18 than is usual here. But undoubtedly the main reason for the apparent reduction in the importance of g in adults is that the testees are more homogeneous in ability." (Vernon 1950,30-31.)

In the 'fifties and 'sixties only few empirical contributions seem to have been published on the age-differentiation issue (e.g., Burt 1954, Cropley 1964), and negligible interest for the problem is apparently the case. However, it might be that a new interest is about to be stirred, prompted by, for example, Guilford (1967), Anastasi (1970), and Reinert (1970), who discuss the matter substantively. Also, there is an increasing concern for methods research in the measurement of change.

Recently, Olsson and Bergman (1973) have involved themselves in the age-differentiation hypothesis, mostly challenged by problems connected with the statistical methodology to be used in testing the hypothesis. But their engagement is probably also reflecting an enduring interest in Sweden for the last decade concerning longitudinal studies on mental development (e.g., Berglund 1965, Ljung 1965, Härnqvist 1968).

Neither in the past nor in the present discussion does it seem to enter much theory in the age-differentiation hypothesis, such that one is able on convincing, rational premises to predict either confirmatory or disconfirmatory evidence on the hypothesis. The problem is therefore as yet to a very great extent a challenge to contriving studies that are noncontroversial as regards the total research plan, comprising issues like samples to be selected, tests to be administered, age range to be covered, and analytical approaches. It seems though almost impossible to include all favorable features of such an investigation in one study. Therefore, a diversity of plans should be encouraged, such that various methods with different weaknesses can be employed in repeated attacks on the same problem. This is in the spirit of multi-operationalism, or convergent operationalism, recommended by Campbell and Fiske (1959), and Webb, Campbell, Schwartz, and Sechrest (1966) among others. Although their context was somewhat different, their recommendation certainly also applies to the present problem.

With a firm belief in a philosophy of research strategy that favors divergent methods on convergent substantive problems, it is the purpose of the present paper to introduce an analysis of variance approach to the testing of the age-differentiation hypothesis; and also to present real-world data that are assumed to be good enough to shed some light on the longstanding substantive problem of whether differentiation of abilities takes place with increasing age.

The problem and some investigational considerations.

The total plan for conducting a study of a problem (like the age-differentiation problem) is here called an investigation, in accordance with Cattell (1966), who describes a total research plan as proceeding through four phases, "from a choice of a theory or a hypothesis, to a choice of a relational system for study, and so devising an experimental design to get data on the relations, and thence to a statistical analysis method" (p.51).

As mentioned above, the age-differentiation hypothesis does not seem to have got a firm theoretical foundation. It is a rather loosely conceived hunch. Nevertheless, researchers involved in the hypothesis should be fairly well agreed on what the implications for a relational system should be from a conception of what the logical structure of the problem is.

The basic requirement to a relational system congenial to the age-differentiation problem is that it should be powerful enough to indicate changes in patterns of abilities, or configurations of abilities, across occasions, or over time. While a simple measurement of change is concerned with the difference between at least two scores for persons at two or more points in time, the age-differentiation problem involves a rather complex conception of a measurement of change in that it is really a question of being concerned with a difference between differences of scores of persons of at least two simultaneous measures on two or more occasions. More concretely, the age-differentiation hypothesis implies that the correlation between two tests (the configuration of the two tests) should change with age, that is, decrease with increasing age. This is a change-in-configuration problem, and assumed to be the heart of the age-differentiation hypothesis.

Ideally, the mathematical system to be employed in the analysis of data gathered to test the age-differentiation hypothesis should be so strong that it can handle conditional relations simultaneously. This means that it should be able to tell whether the correlation between the tests depends on occasion.

The consequences for the design of an experiment on the hypothesis following from the basic conception of what the problem is, as here sketched, would ideally be to have a large, random group of testees take the same test battery on at least two occasions, appropriately separated in time.

The statistical analysis method favored in studies concerning mental organization is undoubtedly factor analysis. As a matter of fact, there seems to have been no other choice. This does not mean, we think, that the method is perfectly fitted for the problem. Rather, the method has obvious drawbacks. According to Reinert (1970) and Olsson and Bergman (1973), the indicators used for testing the age-differentiation hypothesis by way of factor analysis are influenced by quite a few arbitrary decisions, connected with the choice of number of common factors and preferences for methods of rotation. These drawbacks are all associated with the a posteriori character of most factor analysis in use. In addition, it also appears correct to conclude that factor analysis does not meet the ideal requirement of directly being able to test the conditional configuration hypothesis.

Compared to an a posteriori strategy customarily followed in factor analytical approaches to the age-differentiation hypothesis, an a priori strategy would classify (factorize) tests on a rational basis and elaborate on the changing relative contribution made by the categories in this particular classification to total test variance from one occasion to the other.

The categories of tests define the psychologically meaningful factors (a priori linear combinations) to the researcher. Therefore, there is no a posteriori problem with number and types of common factors, and a rotation procedure is unnecessary. Thus, many of the arbitrary decisions in traditional factor analysis approaches can be avoided. This a priori kind of factor analysis is just what an analysis of variance approach will enable us to do.

In view of the particular role played by the general factor in the age-differentiation hypothesis, it might seem natural to prefer a hierarchical factor structure which makes allowance to what is common to all tests in a battery (the g-factor), and also to what is common to major and minor groups of tests (major and minor group factors). This can be accomplished by factor analysis approaches too, but often rotations are made to extract more group factors, deleting the g-factor, rather than being interested in the g-factor per se.

The mathematical system of analysis of variance is on the level of latent structures by way of variance components hierarchically ordered. By choosing a system that generates a hierarchical factor structure, a confirmatory study on the age-differentiation hypothesis should indicate that the composition of test score variance changes from younger to older age groups of youngsters such that relatively more variance can be attributed to major and minor group factors with increasing age and relatively less to the g-factor. It is this potentiality of analysis of variance as a powerful correlational technique that is going to be exploited in the following.

A longitudinal model.

Consider what is generally judged the most preferable design for assessing the age-differentiation hypothesis: A randomly selected group of n persons (P) is tested at time O_1 with a test battery of k tests (T) for each of, say, 2 groups of tests (G); each group of tests (G_1 and G_2) supposed to measure relatively independent traits. Further, assume that the same sample of persons is retested at time O_2 with the same test battery. This will

TABLE 1. Test design for a longitudinal study.

		O_1		O_2	
		G_1	G_2	G_1	G_2
		$T_1 \ldots\ldots T_k$	$T_{k+1} \ldots T_{2k}$	$T_1 \ldots\ldots T_k$	$T_{k+1} \ldots T_{2k}$
	1				
	.				
P	.				
	.				
	n				

constitute a longitudinal study. The test design is presented in TABLE 1.

In analyzing data gathered by this test design, a concern with the age-differentiation hypothesis implies that one is particularly interested in examining the sources of variance that have a P in their labels, because these sources are conveying information on inter- and intraindividual differences.

Therefore, in writing out the analysis of variance (ANOVA) table, only the P sources will be included, such that the O, G, OG, T:G, and TO:G sources are all omitted from TABLE 2. They are the sources of variance connected with differences between tests.

TABLE 2. Structural models for linear combinations.

SS source	df	MS	Structure equations
SS_p	(n-1)	MS_p	$\sigma^2_{pto:g}+2\sigma^2_{pt:g}+k\sigma^2_{pgo}+2k\sigma^2_{pg}+2k\sigma^2_{po}+4k\sigma^2_p$
SS_{po}	(n-1)	MS_{po}	$\sigma^2_{pto:g}+k\sigma^2_{pgo}+2k\sigma^2_{po}$
SS_{pg}	(n-1)	MS_{pg}	$\sigma^2_{pto:g}+k\sigma^2_{pgo}+2k\sigma^2_{pg}$
SS_{pgo}	(n-1)	MS_{pgo}	$\sigma^2_{pto:g}+k\sigma^2_{pgo}$
$SS_{pt:g}$	2(k-1)(n-1)	$MS_{pt:g}$	$\sigma^2_{pto:g}+2\sigma^2_{pt:g}$
$SS_{pto:g}$	2(k-1)(n-1)	$MS_{pto:g}$	$\sigma^2_{pto:g}$

It should be understood that the MS's in TABLE 2 are all variances of linear combinations of the 4k observations in the total test design. The linear combinations should therefore be regarded as factors.

A full-rank solution would mean that 4k orthogonal linear combinations have to be identified in the test design. In effect, this is what has been done in TABLE 2. That is not immediately apparent, however, as the PT and PTO linear combinations are averages of the (k-1) orthogonal linear combinations within each of the 2 groups of tests. That is, for PT:G and PTO:G there are within each of the groups 2(k-1) orthogonal linear combinations of the same category, altogether 4(k-1). Hence, the number of linear combinations included in the SS column are 4 (for P, PO, PG, and PGO) + 2(k-1) for PT:G + 2(k-1) for PTO:G = 4k, which is the number of observations for each person tested. (For a discussion of orthogonal linear combinations and related issues, see, for example, Kendall (1961), Morrison (1967), Hays (1963)).

The meaning of these linear combinations should be made clear: The MS_p is the variance for the sum score of all 4k observations, that is, the sum score of 2k tests across the two occasions. It is a measure of what is common variance to the total test battery.

The MS_{po} is the variance of a linear combination that contrasts the tests from O_1 with the same tests for O_2. This is a measure of the extent to which the two occasions are measuring different things.Otherwise put, it is the variance of the difference score between occasions.

The MS_{pg} is the variance of a linear combination that contrasts the two groups of tests (G_1 versus G_2) for both occasions combined. This is a measure of the extent to which the two groups of tests measure different constructs, irrespective of occasion. Again it is a difference score variance: The variance of the difference score between groups of tests.

The MS_{pgo} is the variance of the linear combination that contrasts two contrasts. As a matter of fact, it compares the difference between the scores of G_1 and G_2 for O_1 with the difference between the scores of G_1 and G_2 for O_2. Now the contrast between the G's is also a measure of the extent to which the G's correlate: Comparing the G's by way of correlation is complementary to comparing the G's by interaction. Hence, it is indeed reasonable to consider the MS_{pgo} in the context of the age-differentiation hypothesis as a source of variance that is at the very heart of the matter: What the MS_{pgo} tells is tantamount to indicating to what degree the correlation between G_1 and G_2 is equal for O_1 and O_2. It is the indicator of the measure of the conditional configuration that one is looking for. It seems correct to regard the MS_{pgo} as the crucial measure of variance to be used in a more direct test of the age-differentiation hypothesis.

Further, the $MS_{pt:g}$ is the pooled variance for linear combinations that concerns the contrasts between tests within the groups for combined occasions. Again it can be viewed as a complement to the correlation among tests within groups: The more PT interaction, the less the correlation between tests within the groups of tests.

Lastly, the $MS_{pto:g}$ is the variance of a linear combination that contrasts the differences between tests within groups for the two occasions. In the context of the present substantive problem, the $MS_{pto:g}$ is measuring the extent to which the correlation between tests within groups differs from O_1 to O_2. Considered this way, it is clear that the $MS_{pto:g}$ is indicating to what extent the age-differentiation hypothesis is tenable: A substantial change in the correlation between tests within groups from one occasion to the other would mean a considerable interaction of the $MS_{pto:g}$ type. The direction of change in correlations among tests would indicate whether the age-differentiation hypothesis is corroborated or falsified.

The MS's of TABLE 2 are all variances of manifest or observed linear combinations of the tests that can be interpreted to convey information of distinct, substantive character. However, observed data generated by fallible measurements will always be looked upon as not quite trustworthy. This means that the observed variances reviewed above should be considered partly inflated by irrelevant variances. Thus in the MS's irrelevant variances are thought to be mixed up with relevant variances of what might be regarded as measures of some underlying, true condition. What is desirable is to get indications of the magnitude of true score variances, or of the relative contribution of true score variances to the fallible, observed score variances.

In order to approach a solution to this problem one has to go beyond the observations made and infer some latent structures on the manifest variances. This is what has been done in the column for structure equations in TABLE 2. In that column the MS's are conceived as being composed of a sum of weighted variance components. The intention is, by inference, to construct a composition of the variances such that some true contribution made by the specified sources of variance can be indicated.

The structure equations in TABLE 2 are all reflecting a theory of how data in this particular test design are generated. For example, the structure equation for MS_{pgo}, the crucial source of variance in the present study, can be interpreted to mean that the observed MS for the PGO interaction is influenced both by a PTO:G interaction, which can be viewed as conveying error variance together with information on the correlation of tests within groups of tests as dependent on occasion, and a PGO interaction proper that reflects the geneuine conditional correlation of groups of tests, i.e., how the correlation of the groups of tests is dependent on occasion. Although the structure is abstract, the equation for MS_{pgo} is real enough in that it's functioning can be illustrated by concretely manipulating the correlations among tests within groups for each occasion such that the correlation between the two groups of tests will automatically be affected.

As no further effort will be made to explicate the basic nature of the latent structures presented in TABLE 2, the reader is referred to a more thorough introduction to and a discussion of variance components analysis. Of particular relevance in the present context/is the conception of a descriptive variance components analysis as presented by the author elsewhere (Eikeland 1971, 1972a, 1972b, 1973).

A latent structure analysis elaborates on the magnitudes of the variance components. By judging the relative contribution by the components to the total variance of the test battery across occasions, one can see how much of the variance is contributed by those components that/are especially relevant for assessing the age-differentiation hypothesis. In the present design those components are σ^2_{pgo} and $\sigma^2_{pto:g}$, particularly the former.

It should be remarked that the author considers a variance components analysis applicable either as a purely descriptive analysis, like factor analysis is used, or as an explanatory analysis in more traditional analysis of variance contexts. In aiming at an explanatory study more emphasis will be put on specifying which of the components should be regarded as signal components, respectively noise components, to be able to indicate the magnitudes of some true score variances. The nature of data to be analyzed will have to decide whether a study should be considered descriptive or explanatory. In the context of the present substantive problem the descriptive variance components analysis will be emphasized.

A cross-sectional model.

A longitudinal model requires repeated measures on the same persons over occasions. This kind of data may be difficult to obtain. In that case, testing different representative samples

TABLE 3. Test design for a cross-sectional study.

	O_1		O_2	
	G_1	G_2	G_1	G_2
	$T_1 \ldots\ldots T_k$	$T_{k+1} \ldots T_{2k}$	$T_1 \ldots\ldots T_k$	$T_{k+1} \ldots T_{2k}$
S_1 (1 . . n_1)			No observations	
S_2 (1 . . n_2)	No observations			

of a target population at different age levels is an alternative choice, commonly considered second best. Such a cross-sectional design will look somewhat different from the longitudinal one. Assuming a test battery of the same design as the one used to describe the longitudinal model, a cross-sectional study is illustrated in TABLE 3 where the symbols are the same as in TABLE 1. In addition, S_1 and S_2 are used for the two samples.

Characteristic for a cross-sectional model as compared to a longitudinal one, is that no cross-occasion covariance matrix will be available. Only two cross-test covariances can be generated, one for each of the two occasions. Therefore, no direct information of change in configuration can be obtained in a cross-sectional design. Only indirectly, by comparing the variance structures for different samples at different age levels, can changes be indicated.

The formal analysis of cross-sectional data will have to be made separately for the data matrices of each sample-occasion combination. The design presented in TABLE 3 means that two data matrices should be analyzed, the S_1O_1 and the S_2O_2 matrices.

TABLE 4. Structural models for linear combinations.

SS sources	df	MS	Structure equations
P	(n-1)	MS_p	$\sigma^2_{pt:g} + k\sigma^2_{pg} + 2k\sigma^2_p$
PG	(n-1)	MS_{pg}	$\sigma^2_{pt:g} + k\sigma^2_{pg}$
PT:G	2(k-1)(n-1)	$MS_{pt:g}$	$\sigma^2_{pt:g}$

Assume that the test battery consists of two groups of tests with k tests within each of the groups, like the battery used for the longitudinal model. The model to be used in the cross-sectional study is specified in TABLE 4. As can be seen it is simpler than the corresponding longitudinal model in TABLE 2. This is because occasion does not enter the cross-sectional model, formally.

As before, only sources of variance that are descriptive of individual differences are included in the ANOVA table. For this specific test design there are three sources of relevance for the analysis, P, PG, and PT:G, while the G and T:G sources are of no interest and omitted from the table.

The P source is descriptive of individual differences in sum score across all of the 2k tests; the PG source of the extent to which groups of tests are measuring different constructs; and the PT:G source of the extent to which the tests within groups are measures of a common construct.

The three sources of variance in TABLE 4 indicate what kind of linear combination is used to account for total test variance in the battery of 2k tests. The linear combination of P is the one so commonly used in differential psychology, the sum score. The combination of PG is the difference score between G_1 and G_2. The PT:G variance is based on the pooled sum of squares for (k-1) orthogonal linear combinations of the k tests within each of the two groups. This means that the PT:G variance includes 2(k-1) linear combinations that contrast the tests within the groups of tests. Thus, the ANOVA has transformed the variance of the 2k tests into the variances of 2 + 2(k-1) orthogonal linear combinations, which add to 2k. The parallel to factor analysis is clear, although it should be noted that the factors here are determined on a rational basis. They are a priori factors.

The equations in TABLE 4 define the latent structures of the variances of the linear combinations. The relative magnitude of the variance components compared for age levels will be of crucial concern in testing the age-differentiation hypothesis, as these components also convey information on how much groups of tests are correlated, and how much tests within groups are correlated.

An increase in the relative contribution to total test variance made by the P component, the σ^2_p, from age level to age level will indicate integration rather than differentiation; while an increase in the relative contribution made by σ^2_{pg} will mean a corroboration of the hypothesis.

A comparison of the composition of the sum score variances, the MS_p's, will according to the structure equation indicate to what extent the sum score variances for age levels are differentially influenced by a general factor, σ^2_p; two major group factors, σ^2_{pg}; and minor group factors, or specific factors, as represented by $\sigma^2_{pt:g}$.

Analysis of longitudinal data.

a) Data

The same data as used by Olsson and Bergman (1973) in their factor analytical approach to the age-differentiation hypothesis will here be reprocessed by the analysis of variance approach as described above. Their sample consists of 260 girls from a Swedish town. The girls were tested in 1965 when 9-10 years old, and retested three years later when they were 12-13 years old. They were given six ability tests and two achievement tests (in Swedish and Mathematics) at both age levels. The tests were not the same on the two occasions, although of supposedly equivalent content.

b) Grouping of tests

The eight tests administered were, in the order presented by Olsson and Bergman (1973): Similarities, Opposites, Swedish, Letter groups, Figure sequences, Mathematics, Cube counting, and Metal folding.

The present approach assumes tests to be grouped a priori. Some rational classification of the eight tests listed above is therefore necessary in order to be able to match the same factors on the two occasions.

It does not seem farfetched and unreasonable to consider making a dichotomy of the eight tests, a dichotomy which according to British factor analysis studies appears repeatedly; the socalled <u>v:ed</u> and <u>k:m</u> factors. Says Vernon (1950, 23): "After removal of g, tests tend to fall into two main groups: the verbal-numerical-educational on the one hand (referred to as v:ed factor), and the practical-mechanical-spatial-physical on the other hand (referred to as k:m factor)." In accordance with these findings, one might group the Similarities, Opposites, Swedish, and Mathematics tests as a v:ed factor. The rest,Letter groups, Figure sequences, Cube counting, and Metal folding might appear to be an approximate match to the k:m factor. As a matter of fact, the plausibility of this particular classification is fairly well supported by the intercorrelation matrix of the eight tests.

c) Analytical procedure

In the analysis to be performed, the two specified groups of tests together with the four tests within each of the groups and the two occasions will define a threefacet test design. For a further discussion of multifacet designs, the reader is referred to Cronbach, Gleser, Nanda, and Rajaratnam (1972), and to Eikeland (1972a).

In keeping with much of traditional factor analysis, the present analysis will elaborate on the correlation matrix. This is done partly because data are most conveniently accessible in the form of a correlation matrix. But the main reason is to show how an analysis of variance approach can be conducted when the point of departure is a correlation matrix (a variance-covariance matrix would also do) and not the basic data matrix.

The implication of the intended procedure is that standardized scores, and not raw scores, will be used as the basic observations to be analyzed. As the data matrix of standard scores is not available, the analysis has to start out with the correlation matrix. This will not in any way affect the results, but some essential modifications in the analytical procedure are necessary, such that the MS's can be given as functions of the correlation matrix.

The correlation matrix for the Swedish data will be a structured 16 x 16 matrix. The structure imposed on the matrix is the test design, which is a 2 groups by 2 occasions by 4 tests

TABLE 5. Structured total intercorrelation matrix for longitudinal data.

			G_1								G_2							
			O_1				O_2				O_1				O_2			
			T_1	T_2	T_3	T_4	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_5	T_6	T_7	T_8
G_1	O_1	T_1							wbb									
		T_2	761		wwb			wbw				bwb				bbb		
		T	723	718														
		T_4	461	491	639													
	O_2	T_1	737	650	620	394												
		T_2	716	680	635	415	769		wwb									
		T_3	702	710	771	534	763	755				bbb				bwb		
		T_4	489	532	660	726	495	540	657									
G_2	O_1	T_5	323	370	501	552	261	260	202	281								wbb
		T_6	310	326	464	505	309	278	251	265	444		wwb					
		T_7	244	291	355	347	358	345	306	358	354	385				wbw		
		T_8	279	304	422	476	528	502	392	483	408	429	461					
	O_2	T_5	361	414	519	514	396	450	532	605	567	465	313	396				
		T_6	319	341	476	459	349	414	461	570	414	605	328	440	603		wwb	
		T_7	258	273	371	484	288	331	346	504	403	367	546	486	444	422		
		T_8	265	283	403	457	283	301	359	539	422	406	432	631	498	532	520	

$\bar{r}_{wwb} = 0,553$ $\bar{r}_{wbw} = 0,658$ $\bar{r}_{wbb} = 0,497$ $\bar{r}_{bwb} = 0,400$ $\bar{r}_{bbb} = 0,362$

within groups design. The splitting up of the total 16 x 16 matrix in a system of submatrices and supermatrices will naturally lead to a specification of different categories of correlation coefficients. These categories can be identified by examining TABLE 5, where the structured intertest correlation matrix is presented. In the upper-right off-diagonal triangular matrix the categories are specified, while the correlation coef-

Insert TABLE 5 about here

ficients as obtained are given in the lower-left off-diagonal triangular matrix.

Altogether 5 categories of correlation coefficients have been defined by taking into account whether the correlation is between or within the three facets of the test design: group of tests, occasion, or test. This is the same classification as a formal analysis of variance approach will come to exploit.

In TABLE 5 the symbols used are G_1 and G_2 for the v:ed and the k:m factors; O_1 and O_2 for the two occasions; T_1- T_4 for Similarities (1), Opposites (2), Swedish (3), and Mathematics (4) tests; T_5- T_8 for Letter groups (5), Figure sequences (6), Cube counting (7), and Metal folding (8). The first subscript for correlation coefficients denotes group, the second occasion, and the third test.A between correlation is denoted b, and a within correlation w.

When a classification of correlation coefficients has been found worth while, it is because the different categories convey distinct information on how various combinations of observations in the test design go together. Some categories of correlation can be expected to be different. Thus, for example, it is reasonable to think that the correlation between groups, between occasions, between tests, $\bar{r}_{bbb}$, will be lower than the correlation within groups, between occasions, within tests, $\bar{r}_{wbw}$. This expectation follows from the assumed reasoning behind the test design.

As is well known, the sum of a correlation matrix is commonly called the test variance, our symbol being V_X, meaning the variance of the sum score. In the present case this will be the variance of the sum score across groups, occasions, and tests; i.e., the variance of the linear combination called P in the model. In defining the MS's for linear combinations PO, PG, and PGO in the models presented in TABLE 2, one has only to change signs for the different combinations. How these signs should run is

TABLE 6. Design matrix for P, PG, PO, and PGO.

	G_1								G_2							
	O_1				O_2				O_1				O_2			
	1	2	3	4	1	2	3	4	5	6	7	8	5	6	7	8
P	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
PG	+	+	+	+	+	+	+	+	–	–	–	–	–	–	–	–
PO	+	+	+	+	–	–	–	–	+	+	+	+	–	–	–	–
PGO	+	+	+	+	–	–	–	–	–	–	–	–	+	+	+	+

shown in the design matrix in TABLE 6, which accounts for only 4 of the 16 orthogonal linear combinations in the total design.

In defining the MS's for PT:G and PTO:G a different procedure has to be followed, as only submatrices of the total correlation matrix should be exploited for these variances. The between groups part of the correlation matrix is of no relevance for that purpose.

The author has elsewhere shown how multifacet analyses can be performed on structured correlation matrices, or variance-covariance matrices. The present design is an extension of the more simple designs discussed by Eikeland (1972b, 1973).

If all coefficients in a correlation matrix are being replaced by the average coefficients of the category to which the particular coefficients belong, the sum score variance for a test design with 2 groups, 2 occasions, and k tests within groups will be,

$$V_X = 4k + 4k(k-1)\bar{r}_{wwb} + 4k\bar{r}_{wbw} + 4k(k-1)\bar{r}_{wbb} + 4k^2\bar{r}_{bwb} + 4k^2\bar{r}_{bbb}$$

It can be shown that the test score variance as commonly defined in differential psychology, i.e., the sum score variance here symbolized V_X, is greater than the corresponding MS_p of the same linear combination by a factor of 4k in the more general case presented above, and by a factor of 2 x 2 x 4 = 16 in the data to be analyzed. Generally, the V_X is greater than the corresponding MS_p by a factor equal to the number of observations going into the test design (Eikeland 1972b). Hence, the MS_p for a test design with 2 groups, 2 occasions, and k tests within groups as it might be computed from the standardized data matrix will be in terms of the properties of the correlation matrix,

$$MS_p = 1 + (k-1)\bar{r}_{wwb} + \bar{r}_{wbw} + (k-1)\bar{r}_{wbb} + k\bar{r}_{bwb} + k\bar{r}_{bbb}$$

The relationship established between V_X and MS_p will hold for other linear combinations as well. The MS's for all of the six sources of variance for individual differences in the model of TABLE 2 can thus be worked out for the empirical data as shown in TABLE 7. It should be noted that the MS's for the nested

TABLE 7. Definitions of MS's for linear combinations.

P	$1 + 3\bar{r}_{wwb} + \bar{r}_{wbw} + 3\bar{r}_{wbb} + 4\bar{r}_{bwb} + 4\bar{r}_{bbb}$
PG	$1 + 3\bar{r}_{wwb} + \bar{r}_{wbw} + 3\bar{r}_{wbb} - 4\bar{r}_{bwb} - 4\bar{r}_{bbb}$
PO	$1 + 3\bar{r}_{wwb} - \bar{r}_{wbw} - 3\bar{r}_{wbb} + 4\bar{r}_{bwb} - 4\bar{r}_{bbb}$
PGO	$1 + 3\bar{r}_{wwb} - \bar{r}_{wbw} - 3\bar{r}_{wbb} - 4\bar{r}_{bwb} + 4\bar{r}_{bbb}$
PT:G	$1 - \bar{r}_{wwb} + \bar{r}_{wbw} - \bar{r}_{wbb}$
PTO:G	$1 - \bar{r}_{wwb} - \bar{r}_{wbw} + \bar{r}_{wbb}$

linear combinations, PT:G and PTO:G, have to be written as functions of the submatrices for G_1 and G_2. No between groups correlation can enter those MS's.

As mentioned in the discussion of the basic model for a longitudinal study, TABLE 2, the total variance of the 4k observations going into the test design will in the analysis of variance approach be transformed into a sum of the variances of the 4k orthogonal linear combinations defined by the particular design constructed. Of those 4k combinations, the PT:G and PTO:G will each have 2(k-1) linear combinations within themselves. The MS's for PT:G and PTO:G are thus average values of the variances of the 2(k-1) linear combinations.

If the MS's for PT:G and PTO:G in TABLE 7 are each of them multiplied by $2(k-1) = 6$, and the whole table added, one obtains a sum of 16. For those of the readers who are familiar with a full rank factor analysis solution, it should be clear that the sum has to add to 16, which is the trace of the correlation matrix, i.e., the sum of the principal diagonal in the matrix.

The MS's in TABLE 7 are variances of manifest linear combinations of the observations made. As we want to go beyond the manifest level to an inferred latent variance structure, we need to find a way to defining some true contribution made by the sources of variance in the test design. These latent trait contributions to total test variance are represented by the components.

Even by only having access to the correlation matrix, the variance components can be found. This is accomplished by solving for the unknowns in the structure equations in TABLE 2 (recalling that $k = 4$) by starting from the bottom of the table.

TABLE 8. Definition of variance components.

σ^2_{p}	$\bar{r}_{bbb}$
σ^2_{pg}	$\bar{r}_{wbb} - \bar{r}_{bbb}$
σ^2_{po}	$\bar{r}_{bwb} - \bar{r}_{bbb}$
σ^2_{pgo}	$\bar{r}_{wwb} - \bar{r}_{wbb} - \bar{r}_{bwb} + \bar{r}_{bbb}$
$\sigma^2_{pt:g}$	$\bar{r}_{wbw} - \bar{r}_{wbb}$
$\sigma^2_{pto:g}$	$1 - \bar{r}_{wwb} - \bar{r}_{wbw} + \bar{r}_{wbb}$

The definitions of the variance components are given in TABLE 8. It is indeed interesting to see how the components can be defined as functions of average correlation coefficients of the correlation matrix.

As the six components in TABLE 8 are inferred to be additive elements of the variance of one observation, which in the correlation matrix is 1, it is not unexpected that the sum of the defined components should add to 1.

d) Results

The analysis of variance of the correlation matrix is conducted in accordance with the definitions of MS's and variance components given in TABLE 7 and TABLE 8. In keeping with the structure equations of TABLE 2, the latent variance structures of the MS'S are also given.

TABLE 9. Analysis of variance results.

		Latent variance structure	Variance
Source	MS	pto:g pt:g pgo po pg p	components
P	7,856	= 0,286+0,322+0,072+0,304+1,080+5,792	0,362
PG	1,760	= 0,286+0,322+0,072+1,080	0,135
PO	0,662	= 0,286+0,072+0,304	0,038
PGO	0,358	= 0,286+0,072	0,018
PT:G	0,608	= 0,286+0,322	0,161
PTO:G	0,286		0,286
		Sum =	1,000

As pointed out in the discussion of the model, it is the PGO interaction that is of most concern in testing the age-differentiation hypothesis, in that this particular linear combination conveys information on how much of the variance in the system is generated by a change in the correlation between the groups of tests from the first testing at age 9-10 to the second testing three years later. In examining the variance structure of one observation, it is evident from the column of variance components that almost no variance can be inferred to be contributed by σ^2_{pgo}, less than 2%, which indicates that negligible change in the configuration of the groups of tests has taken place from the first to the second testing.

Of the observed variance of 0,358 for linear combination PGO, only 0,072, or 20% of observed variance, is considered the contribution by a true change-in-configuration component. Another way of looking at that result would be to regard the difference between differences of the group scores for the two occasions as rather unreliable. The reliability of the linear combination of PGO can be given the traditional form of taking the ratio of true score variance to observed score variance,

$$r_{tt} = \frac{0,072}{0,286 + 0,072} = 0,201,$$

which certainly indicates that the dependability of that difference score is very low. The rationale for difference score reliability is discussed in more detail by Eikeland (1973).

Still another way of interpreting the result is to look at the observed sum score variance, the MS_p, and examine the latent variance structure to learn how much of that variance can be considered explained by a true change in correlation between the two groups of tests from O_1 to O_2. By transforming MS_p to unit variance, the contribution made by the weighted components can be given as proportions,

$$\begin{array}{ccccccccccccc} & & & \text{pto:g} & & \text{pt:g} & & \text{pgo} & & \text{po} & & \text{pg} & & \text{p} \\ MS_p & = & 1{,}000 = & 0{,}036 & + & 0{,}041 & + & 0{,}009 & + & 0{,}039 & + & 0{,}137 & + & 0{,}737 \end{array}$$

The interpretation of this composition of the sum score variance is fairly clearcut: If all observations in the test design, altogether 16, were added, the variance of that sum score could be regarded as strongly loaded by a general component running through all of the observations, the weighted σ^2_p component; only moderately loaded by a weighted component signifying that the two groups of tests are measuring different constructs, the σ^2_{pg} component; and negligibly loaded by other components. Among these, less than 1% is contributed by the component for the PGO interaction, σ^2_{pgo}.

The result commented upon till now has been considered in terms of the latent structure by inferring from a theoretical structure imposed on data. However, one can also look quite unsophisticatedly at data by considering the observed variances of the 16 orthogonal linear combinations as created by the test

TABLE 10. Contribution of MS's to total test variance.

		%
MS_p	7,856	0,491
MS_{pg}	1,760	0,110
MS_{po}	0,662	0,041
MS_{pgo}	0,358	0,022
$6(MS_{pt:g})$	3,648	0,228
$6(MS_{pto:g})$	1,716	0,107
Tot test var	16,000	0,999

design. This will be a parallel to a full rank data reduction analysis. In TABLE 10 the MS's for PT:G and PTO:G are each multiplied by 6 and added together with the MS's for P,PG,PO, and PGO to yield the total variance of all of the 16 linear combinations. As expected, this sum is 16. Thus, the total test variance has been accounted for by the 16 a priori orthogonal linear combinations, constructed on a rational basis to answer specific substantive questions of interest to the researcher. The results presented in TABLE 9 and TABLE 10 are not comparable in principle, as the two analyses are based on quite different ways of conceiving variance structures. In TABLE 9 the structure is latent, in TABLE 10 it is a manifest structure. Nevertheless, the contributions made to total test variance by the sources are so clear that the two ways of analyzing data lead to the same conclusion as regards the problem at issue: The change in correlation between the two groups of tests, the v:ed and the k:m factors, seems to change negligibly from test to retesting three years later.

Analysis of cross-sectional data.

a) Data

Sample test data for the cross-sectional study are taken from the files of the standardizing material for a Norwegian mental maturity test battery gathered as far back as 1952. (For a general discussion of the test battery, see Sandven (1962).)

The battery is a Thurstone type of intelligence tests, covering five primary mental abilities: Memory, Verbal, Spatial, Reasoning, and Quantity. The battery is hierarchically stratified, like PMA and WISC, in that there are subtests (S) nested within tests or abilities (T), and items (I) nested within the subtests.

Series III of the test battery is used. This series is standardized for the age group 12-15. In the analysis to be performed, the number of items in the battery has been cut from originally 114 to 65 by omitting items in the second half of the subtests. This reduction has been done in order to avoid a spuriously high internal consistency because of too many unattempted items. TABLE 11 shows the tests represented in the

TABLE 11. An overview of the test battery.

Test	Subtest	No.of items
Memory	1	14
	2	5
Verbal	3	6
	4	5
Spatial	5	4
	6	4
	7	4
Reasoning	8	6
	9	5
Quantity	10	6
	11	6
		N_i = 65

battery, the listing of subtests from 1 to 11, and the number of successive items retained in the analysis, always starting with item No. 1 in each original subtest.

Samples are randomly drawn from the files of the age group 12-15, the target population being the children of that age range in the three biggest cities in Norway; Bergen, Oslo, and Trondheim. This particular age range was chosen in order to have just the same test battery, Series III, applied to all samples included in the study. The battery had been administered in the fall of 1952 (September) in grades 6 and 7 in elementary school, and in grammar school (realskulen) and the continuation school (framhaldsskulen), which are here called grade 8.

Children born 1940 were in the fall of 1952 6-graders. Those tested were divided into two groups; children born in the first half of 1940, and children born in the second half. This was done for each sexseparately. From each of these four groups of 6-graders, 50 children were randomly picked. The same procedure was followed for children born in 1939; that is, children in grade 7.

Children born in 1938 were in the fall of 1952 either in grammar school or in the continuation school. These schools were selective schools, meaning, generally speaking, that high and medium achievers went to grammar school, while medium and low achievers went to the continuation school. In the three cities concerned, some 60% of the age group went to the former, some 40% to the latter type of school this particular year. Accordingly, in order to get a representative sample for the age group considered, children born in 1938, proportionate samples had to be drawn from the files of each of the two selective school types.

Again, children within grammar school and the continuation school were divided into two groups, those born in the first half of 1938 and those born in the second half.

An assumedly representative sample is now obtained by randomly picking 30 children from each of the age groups in grammar school and 20 from each of the age groups in the continuation school. This sampling procedure is performed for each sex separately.

Thus, by the sampling plan described, representative samples of 50 children are supposed to have been established for six age levels with approximate average age 12.0, 12.5, 13.0, 13.5, 14.0, and 14.5. Altogether 600 children go into the 12 samples of the study.

b) The problem restated

The intention now is to examine how the variance structure of the test battery for the different samples will come to change from age level to age level. For the battery used in this study, it is reasonable, in keeping with the age-differentiation hypothesis, to expect that the correlations among tests (abilities) should decrease with increasing age.

As noted in the discussion of the cross-sectional model to be applied, no direct testing of change is possible. Variance structures for age levels have to be compared descriptively.

c) Modification of the basic model

The basic model to be applied in the cross-sectional study presented in TABLE 4 is based on a twofacet, hierarchical test design, meaning that tests are nested within groups of tests. As the design of the test battery that provides data for the present analysis is a threefacet, hierarchically stratified test design, the model needs an extension to fit this particular structure of tests. The design for this kind of test battery has been discussed more thoroughly by the author elsewhere (Eikeland 1972a), and will not be dealt with in any detail here.

An analysis of variance model for the test battery in question will have to take especially into account that there are three facets, and that the design is doubly nested.

There are four sources of variance that are descriptive of inter- and intraindividual differences in the present test

TABLE 12. Structure equations for the threefacet test design.

SS source	df	MS	Latent structure
SS_p	$(n-1)$	MS_p	$\sigma^2_{pi:s:t}+k\sigma^2_{ps:t}+km\sigma^2_{pt}+kmr\sigma^2_p$
SS_{pt}	$(n-1)(r-1)$	MS_{pt}	$\sigma^2_{pi:s:t}+k\sigma^2_{ps:t}+km\sigma^2_{pt}$
$SS_{ps:t}$	$(n-1)(m-1)r$	$MS_{ps:t}$	$\sigma^2_{pi:s:t}+k\sigma^2_{ps:t}$
$SS_{pi:s:t}$	$(n-1)(k-1)mr$	$MS_{pi:s:t}$	$\sigma^2_{pi:s:t}$

design: P, PT, PS:T, and PI:S:T. The P source is the linear combination that concerns the sum score across tests, subtests, and items. The PT involves the correlation among tests(abilities), the PS:T the correlation between subtests within tests, and the PI:S:T the correlation among items within subtests within tests. As items within subtests are supposed to measure the same trait, one can also regard PI:S:T as a measure of the inconsistency of items in measuring a common trait within subtests.

In TABLE 12 the structure equations for the extended model are given, assuming k items within subtests (S), m subtests within tests (T), r tests, and n persons (P) tested. As before, only sources of variance involving individual differences are included in the table, which means that only linear combinations of observations for persons are of concern.

It should be noted that the model presented in TABLE 12 is slightly simplified compared to a model that would be a more exact fit to the test battery employed. The simplification concerns the coefficients in the structure equations. As is evident from TABLE 11, there is not the same number of items within subtests, neither is the number of subtests within tests equal. This should not bother us. For the purpose of the present analysis we need not go into the estimation of unweighted components. Only the values of weighted variance components will be used for obtaining the variance structures demanded in order to be able to compare age levels. As can be seen from TABLE 12, when the unknowns are weighted components, no difficulty will arise in solving for them, even if the coefficients were some kind of average number of items and average number of subtests. If unweighted components are wanted, there are formulas for determining the coefficients in case of unequal items within subtests and/or unequal subtests within tests.

d) Results

Altogether 12 data matrices, each of them a 50 persons by 65 items matrix have been analyzed. The data processing had to be done manually as no computer program was available.

Only one of the 12 analyses will be presented here in order to show how results were worked out. In TABLE 13 a complete

TABLE 13. ANOVA of data matrix for sample B138.

Source	SS	df	MS	Variance structure pi:s:t	ps:t	pt	p
P	48,235	49	0,9844	=0,1537	+0,0511	+0,0517	+0,7279
T	14,736	4					
S:T	18,561	6					
I:S:T	54,816	54					
PT	50,265	196	0,2565	=0,1537	+0,0511	+0,0517	
PS:T	60,215	294	0,2048	=0,1537	+0,0511		
PI:S:T	406,605	2646	0,1537				
Total	653,435	3249					

ANOVA table is presented. It is the analysis of the data matrix for the sample of boys of average age 14,5; that is, boys born in the first half of 1938.

In assessing the tenability of the age-differentiation hypothesis, the variance structure for MS_p will be the focus of attention, because that structure includes all of the crucial information on how parts of the test battery go together. By setting MS_p to unit variance for all samples, the proportion of variance accounted for by the weighted components in the structure equation for MS_p can be compared for the different age levels. For the result in TABLE 13 the proportions will be,

$$\begin{array}{lccccc} & & \text{pi:s:t} & \text{ps:t} & \text{pt} & \text{p} \\ MS_p = 1{,}000 = & & 0{,}156 + & 0{,}052 + & 0{,}053 + & 0{,}739. \end{array}$$

Interpreting the obtained composition of weighted variance components, indicates that a common trait running through the whole test battery explains 74% of the observed interindividual differences, 5% is explained by a differential contribution by tests, i.e., a measure of the extent to which tests measure different traits. Another 5% of the sum score variance is explained by the fact that different subtests within tests measure different traits. Lastly, 16% of the sum score variance is accounted for by the inconsistency of items within subtests in measuring a common trait for each subtest.

In the context of the present study the variance structures per se are hardly of any concern. Rather, it is the variance structures compared for the age groups that count.

When analyses like that illustrated in TABLE 13 have been conducted for the other data matrices, each MS_p is set to unit variance such that all of the variance structures can be given as proportions, and the compositions can be directly compared. It is the relative contribution of the weighted components to MS_p that is relevant for the age-differentiation hypothesis, especially the contribution made by the PT component. If the relative contribution of that component increases with age at the cost of the P component, then differentiation is indicated.

The variance structures of MS_p for all 12 samples are presented in TABLE 14. The samples have been identified by labels

TABLE 14. The variance structure of the MS_p for all samples.

Sample	pi:s:t		ps:t		pt		p
6G240	0,152	+	0,070	+	0,054	+	0,724
6G140	0,128	+	0,050	+	0,061	+	0,761
7G239	0,151	+	0,059	+	0,037	+	0,753
7G139	0,183	+	0,101	+	0,023	+	0,692
8G238	0,139	+	0,034	+	0,131	+	0,696
8G138	0,127	+	0,078	+	0,007	+	0,788
Mean for G	0,147	+	0,065	+	0,052	+	0,736
6B240	0,178	+	0,055	+	0,108	+	0,659
6B140	0,134	+	0,042	+	0,119	+	0,723
7B239	0,143	+	0,064	+	0,089	+	0,704
7B139	0,152	+	0,050	+	0,088	+	0,710
8B238	0,138	+	0,084	+	0,006	+	0,772
8B138	0,156	+	0,052	+	0,053	+	0,739
Mean for B	0,150	+	0,058	+	0,077	+	0,718

indicating grade, sex, and when born. For example, 8G238 means the sample of girls in grade 8 (either in grammar or continuation school) born in the second half of the year 1938; 7B139 the sample of boys in grade 7 born in the first half of 1939.

Figure 1 is drawn to facilitate the interpretation of TABLE 14 with a view to the substantive problem raised of whether a

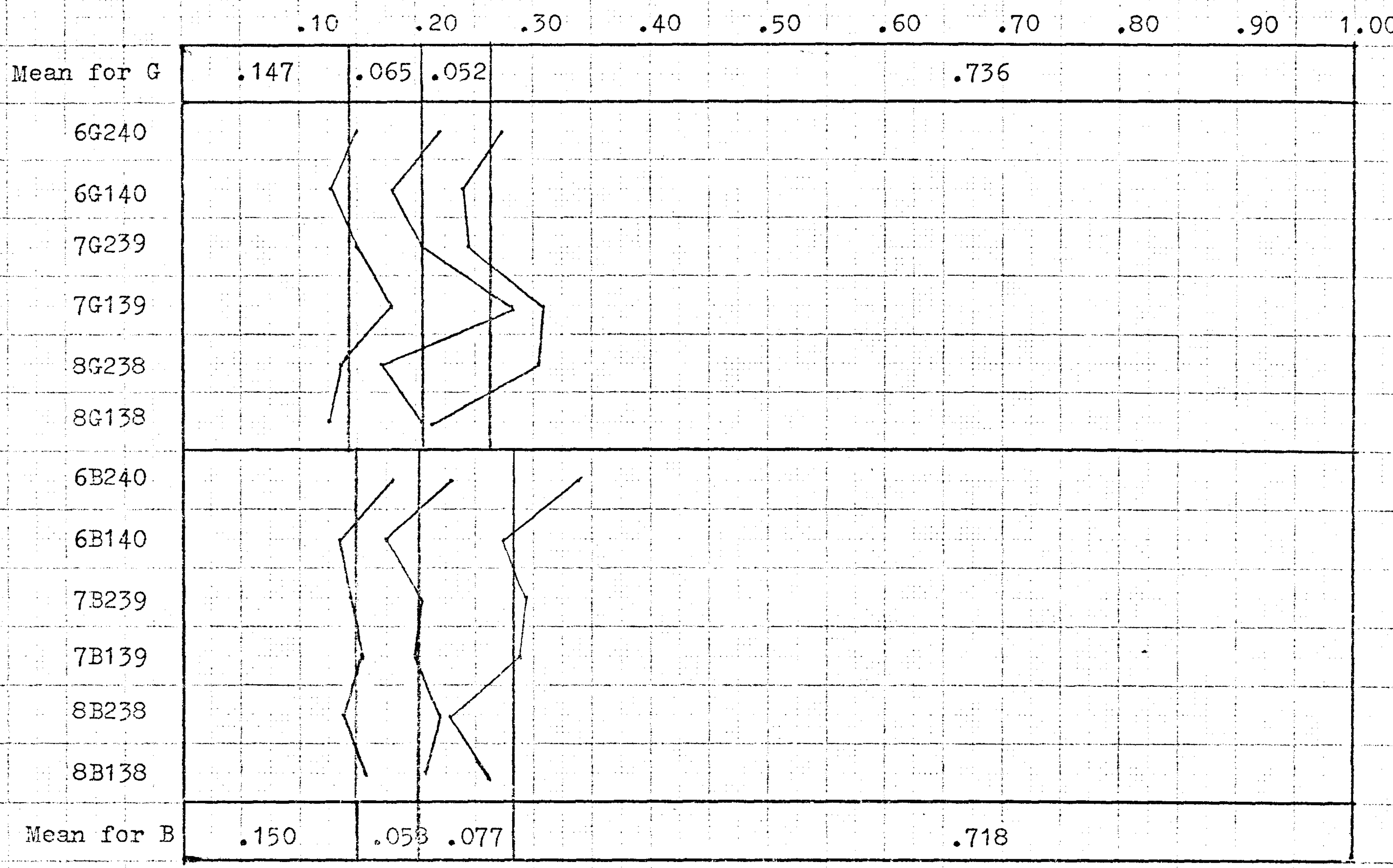

FIGURE 1. The structure of sum score variance compared for samples.

change in variance structure can be substantiated as a function of age. It is obvious from the figure that no change in the relative contribution by the different variance components has occurred from age level 12,0 to age level 14,5 for either sex.

Insert FIGURE 1 about here

It seems reasonable to consider the observed differences in variance structure from age level to age level as unsystematic fluctuations about the means for the two sexes. No trend can be discerned,neither for differentiation nor for integration.

Discussion.

a) The longitudinal study

In the model developed for the longitudinal study, the one most important source of variance for testing the age-differentiation hypothesis is a linear combination of observations, a difference-of-differences score, that signifies a triple interaction, the PGO. For a long time in the history of analysis of variance such interactions were considered bothersome to the researcher, of doubtful substantive interest, and difficult to interpret. In recent years, however, interactions have caught the attention of research workers more and more as they see that realistic, complex questions can be put this way.

In a correlational context, there seems to be great possibilities for applying analysis of variance approaches as an alternative to traditional factor analysis in studies where it can be determined on a logical basis what kind of repeated measures design is appropriate for asking specific questions of data in order to answer particular research questions. It is the author's view that analysis of variance is a superb analytical tool for what can be meaningfully called a rationalistic factor analysis. Guttman's (1958) somewhat vague, but intuitively correct anticipation of what lies ahead for factor analysis,conceived in the frame of reference of analysis of variance, is about to come true. Even Rozeboom (1966) who seems to question a priori factor patterns, must though admit that there might be strong indications in favor of a rationalistic approach:

> "To an empiricist, there is too much about rationalistic analysis that reeks of myth making, too much of coercing nature into strained compliance with our preconvictions when we should be standing aloof with keen but coldly critical vision to register what may be revealed unto us. Yet if the data do, in fact, neatly fit an a priori pattern, then this remains a brute statistical fact about the observed interrelations, untarnished by whatever implausibility may adhere to the pattern as a hypothesis about source variables, which demands due recognition in whatever substantive theory may ultimately be developed about these variables." (Rozeboom 1966, 291.)

The triple interaction in the present model is tailored for asking just that question of data that is implied in the age-differentiation hypothesis. No difficulty attaches to the interpretation of this interaction, although the relationship revealed by it is a complex one. Many research questions that ought to be asked are certainly of a complex character, and should of course not be avoided if one wants to be true to real world relations, which doubtlessly very often are conditional: The relation sought between two variables are frequently dependent on levels of a third variable, and even on the levels of more variables.

The PGO interaction in the model involves a conditional relation in that it delves into the problem of whether the correlation between G_1 and G_2 is dependent on occasion. The small con-

TABLE 15. Changes in correlations between G_1 and G_2 tests from O_1 to O_2.

Tests	O_1	O_2	D_{O_1/O_2}
1/5	0,323	0,396	-0,073
1/6	0,310	0,349	-0,039
1/7	0,244	0,288	-0,044
1/8	0,279	0,283	-0,004
2/5	0,370	0,450	-0,080
2/6	0,326	0,414	-0,088
2/7	0,291	0,331	-0,040
2/8	0,304	0,301	+0,003
3/5	0,501	0,532	-0,031
3/6	0,464	0,461	+0,003
3/7	0,355	0,346	+0,009
3/8	0,422	0,359	+0,063
4/5	0,552	0,605	-0,053
4/6	0,505	0,570	-0,065
4/7	0,347	0,504	-0,157
4/8	0,476	0,539	-0,063
		Average change =	-0,041

tribution made by the PGO interaction to total test variance in the present study is interpreted to mean that the correlation between the G's does not change much from O_1 to O_2.

A cery concrete illustration of the meaning of the PGO interaction is to check directly how the correlation coefficients between tests of the two groups change over time. In TABLE 15 this is shown for all of the 16 bwb correlations. The changes in the coefficients are rather small, but fairly consistent. The average change in the correlation between tests from the two groups of tests from O_1 to O_2 is -0,041, which means that the correlations have increased and not decreased, as predicted by the age-differentiation hypothesis.

Also the PTO:G interaction is thought to be of some pertinence to the age differentiation hypothesis in that a decrease

TABLE 16. Changes in correlations between tests within groups from O_1 to O_2.

Tests	O_1	O_2	D_{O_1/O_2}
1/2	0,761	0,769	-0,008
1/3	0,723	0,763	-0,040
1/4	0,461	0,495	-0,034
2/3	0,718	0,755	-0,037
2/4	0,491	0,540	-0,049
3/4	0,639	0,657	-0,018
5/6	0,444	0,603	-0,159
5/7	0,354	0,444	-0,090
5/8	0,408	0,498	-0,090
6/7	0,385	0,422	-0,037
6/8	0,429	0,532	-0,103
7/8	0,461	0,520	-0,059
		Average change =	-0,060

in the correlation among tests within groups from O_1 to O_1 may indicate differentiation as well. This type of interaction can be illustrated by comparing the correlation coefficients among tests within the two groups for the two occasions. This is shown in TABLE 16, where the changes are consistently in the direction of increasing correlations with age, rather than decreasing as expected from the hypothesis. It is this change that is reflected in the $MS_{pto:g}$.

It should be noted that the direction of changes in the correlation coefficients, whether for PGO or PTO:G, can not be seen from the MS_{pgo} or the $MS_{pto:g}$. It has to be observed from the correlation matrix.

As mentioned in the description of the test battery, the tests used were not the same for the two occasions. To what extent this can have affected the result is difficult to judge.

The present results are based on an analysis of the correlation matrix, or the data matrix of standardized scores. The analysis performed by Olsson and Bergman (1973) of the same data was based on raw scores. It is not believed to have affected the results in the two studies differentially.

b) The cross-sectional study

The formal model used in the cross-sectional study is not so complex as the model developed for the longitudinal study in that occasion can not be directly included in the analysis in the former. The crucial information in the data set generated by the test battery that is used for the assessment of the age-differentiation hypothesis in the cross-sectional study is conveyed first and foremost by the PT interaction. This information concerns the correlation among tests (abilities). How this correlation changes with age is also here the central issue of the age-differentiation hypothesis. As a matter of fact, this way of posing the problem implies a PTO interaction, although this interaction can not be formally specified in the cross-sectional model because no covariance matrix between occasions exists.

Also the cross-sectional model has a clear parallel to factor analysis. In effect, the analysis of the test battery for each age level concerns the factor composition of total test variance. The comparison between age levels will be a study of the invariance of factor structures of the battery over time, as reflected in the relative contribution to total test variance by a general component (factor) and major and minor group components (factors represented by two types of interactions, the PT and the PS:T).

Some doubt may be raised whether the variance structures compared are influenced in an uncontrolled manner by the possibility that the same test battery used for all age levels can be too difficult for the younger children and too easy for the older ones in the age range included in the study. In case of a substantial variation in the dispersion of scores for the different age groups, it is very difficult to know how this will affect the variance structure. If the difficulty level of the test battery is not extreme in either direction for the age levels included in the study, it should be reasonable to think that the structures will not be influenced in any systematic way.

A check on the difficulty level of the test battery together with the mean and the standard deviation of the sum score is given in TABLE 17. As can be seen from that table, the difficulty level of the reduced test battery ranges from 53 (53 items

TABLE 17. Difficulty level, average sum score, and standard deviation for test battery.

Sample	Difficulty	Mean	SD
6G240	54	34,82	8,83
6G140	53	34,38	9,53
7G239	64	41,42	8,60
7G139	62	40,24	7,78
8G238	74	47,80	8,17
8G139	74	48,30	8,51
6B240	62	40,32	8,11
6B140	56	36,08	9,30
7B239	62	40,42	8,84
7B139	61	39,36	8,78
8B238	74	48,18	8,31
8B139	72	46,88	8,00

out of 100 correctly answered) for one of the youngest groups to 74 for the older groups. After all, this does not seem unexpectedly much, and the discrimination power of the test battery should be considered pretty good also for the older groups. To be sure, in planning a study with the same test battery administered to different age groups, an effect like the present one is foreseen, yet unavoidable.

We think it is a sound judgment to consider the result of TABLE 17 in no way invalidating the conclusion reached that no change in the configuration of abilities has taken place from age level 12,0 to age level 14,5 for either sex.

Concluding remarks.

The age-differentiation hypothesis has been approached by analysis of variance models in a longitudinal and a cross-sectional study. As an alternative to a factor analysis approach to the substantive problem of concern in that hypothesis, we think ANOVA has some obvious advantages. These are associated with the a priori character of ANOVA, as contrasted with the a posteriori character of traditional factor analysis approaches. The problem under investigation demands that specific questions should be put to data in order to get unambiguous answers. To obtain this, rational designs in accordance with the problems raised have to be constructed. By so doing, much of the arbitrary decision-making in the factor analysis approach is avoided. Thus, we may be said to have reached a rationalistic factor analysis approach, which is a proper name for the analysis of variance applied in the present context.

Neither of the substantive results obtained in the two studies can be interpreted to corroborate the age-differentiation hypothesis. Rather, at least the result of the longitudinal analysis seems to be more in favor of integration than differentiation. This is somewhat at odds with the result of Olsson and Bergman (1973) by their model 4, which they apparently favor, although they can also see indications in data in the direction of integration. We think those indications are very consistent, as shown in TABLE 15 and TABLE 16.

The cross-sectional study appears to indicate neither differentiation nor integration in the mental organization for the age range included.

References.

Anastasi, A. 1970. On the formation of psychological traits. American Psychologist, 25, 899-900.

Asch, S. 1936. A study of change in mental organization. Archives of Psychology, No. 195.

Berglund, G. 1965. Mental growth: A study of changes in test ability between the ages of nine and sixteen years. Uppsala: Almqvist och Wiksell.

Burt, C. 1954. The differentiation of intellectual ability. British Journal of Educational Psychology, 24, 76-90.

Campbell, D. T. and Fiske, D. W. 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.

Cattell, R. B. 1966. The principles of experimental design and analysis in relation to theory building. Chapter 2 in Cattell, R. B.(Editor). Handbook of multivariate experimental psychology. Chicago: Rand McNally.

Clark, P.M. 1944. Changes in primary mental abilities with age. Archives of Psychology, No. 291.

Cronbach, L. J., Gleser, G. C., Nanda, H., and Rajaratnam, N. 1972. The dependability of behavioral measurement: Theory of generalizability for scores and profiles. New York: John Wiley.

Cropley, A. J. 1964. Differentiation of abilities. Socioeconomic status and the WISC. Journal of Consulting Psychology, 28, 512-517.

Eikeland, H. M. 1971. Correlational analyses of school marks influenced by multiple sources of variance. Explorations into internal structures of complex systems of variation. Oslo:Institute for Educational Research. Mimeographed.

Eikeland, H. M. 1972a. The structure of generalizability theory for hierarchically stratified tests. Oslo: Institute for Educational Research. Mimeographed.

Eikeland, H. M. 1972b. Toward a rationalistic factor analysis. Explorations into a priori covariance structures. Oslo: Institute for Educational Research. Mimeographed.

Eikeland, H. M. 1973. Generalizability estimates for difference scores: An aspect of the construct validity of tests. Oslo: Institute for Educational Research. Mimeographed.

Garrett, H. E. 1938. Differentiable mental traits. Psychological Records, 2, 259-298.

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

Garrett, H. E., Bryan, A. I., and Perl, R. 1935. The age factor in mental organization. Archives of Psychology, No. 176.

Guilford, J. P. 1967. The nature of human intelligence. New York: McGraw-Hill.

Guttman, L. 1958. What lies ahead for factor analysis? Educational and Psychological Measurement, 18, 497-515.

Härnqvist, K. 1968. Relative change in intelligence from 13 to 18. Scandinavian Journal of Psychology, 9, 50-82.

Hays, W. L. 1963. Statistics. New York:Holt,Rinehart & Winston.

Kendall, M. G. 1961. A course in multivariate analysis. New York: Hafner Publishing Company.

Ljung, B. O. The adolescent spurt in mental growth. Stockholm: Almqvist och Wiksell.

Morrison, D. F. 1967. Multivariate statistical methods. New York: McGraw-Hill.

Olsson, U. and Bergman, L. R. 1973. A structural model for testing the age-differentiation hypothesis. Uppsala: Department of Statistics. Mimeographed.

Reichard, S. 1944. Mental organization and age level. Archives of Psychology, No. 295.

Reinert, G. 1970. Comparative factor analysis of intelligence throughout the human life-apan. In Goulet, L. R. and Baltes, P. B. (Editors). Life-span development psychology, 467-484. New York: Academic Press.

Rozeboom, W. W. 1966. Foundations of the theory of prediction. Homewood, Illinois: The Dorsey Press.

Sandven, J. 1962. Det teoretiske og metodiske grunnlag for modenhetsprøving. Pedagogisk Forskning (Scandinavian Journal of Educational Research), 7, 147-168.

Vernon, P. E. 1950. The structure of human abilities. London: Methuen and Co. Ltd.

Webb, E. J., Campbell, D. T., Schwartz, R. D., and Sechrest, L. 1966. Unobtrusive measures: Nonreactive research in the social sciences. Chicago: Rand McNally.