13

TOWARD A RATIONALISTIC FACTOR ANALYSIS

Explorations into a priori covariance structures

Hans-Magne Eikeland

University of Oslo

Oslo, April 1972

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.

PREFACE

An enduring preoccupation for quite a few years now with analysis of variance (ANOVA) as a descriptive data-analytic technique, has lead to a firm conviction that ANOVA should be looked at independent of the traditional linkage to experimental design and strict probabilistic (statistical) thinking.

When at long last one has reached what seems a kind of fresh insight that ANOVA can be used as a most powerful correlational technique, then great possibilities in data analysis seem to open up in borderline cases where traditional taxonomies of techniques could provide no guidelines for analysis.

The author experienced this in dealing with the convergence of ANOVA on multiple regression (Eikeland 1971). The same feeling is prevailing in dealing with the convergence of ANOVA on FA (factor analysis).

While one feels happy on a conceptual level, the feelings are mixed up when it comes to the presentation of the convergence of concepts from different traditions. This is so because one has to compromise as regards exact formulations, use of symbols,etc., in an effort to make things understandable and meaningful to people in the applied research field.

The compromise reached as to mode of presentation in this monograph,is based on the author's experience of how difficult it is to come across with the rationales for complex data analysis techniques to colleagues that are not trained in the formal aspects. Their reactions will show to what extent one has been successful.

Thanks are due to my friends and colleagues, Ola O. Bø, Torleif Lund, and Finn Tschudi, who encouraged and prompted me to communicate to others in written form on these issues.

Oslo, April 1972.

Hans-Magne Eikeland

CONTENTS

1. INTRODUCTION

Analysis of variance (ANOVA), multiple regression analysis (MR), and factor analysis (FA)[1], are by most researchers in psychology and education regarded as nonoverlapping techniques, having their own rationales and distinct functions in the analysis of data. However, the formal system and the basic logic underlying these techniques, and even others not mentioned here, are the same. Mathematicians call this common foundation the general linear model. (See Fennessey (1968) as a readable reference.)

---

1) Factor analysis (FA) will in this report be used as a generic term, covering both data reduction techniques, like the principal components analysis and the centroid solution with 1's in the principal diagonal of the correlation matrix, and classical factor analysis techniques with estimates of communalities in the principal diagonal.

Excellent taxonomies of statistical techniques can be found in the literature. Notable among these is Tatsuoka & Tiedeman (1963), where techniques are classified according to the role, scale type, and number of variables involved. Certainly, such taxonomies are extremely useful. However, when discriminating features are emphasized and convergent features ignored, too rigid classifications of techniques are established, such that possibilities of teasing out information in data are not exploited.

Fortunately, efforts have been made, and more and more efforts are being made, to make research workers aware of the relations among the techniques in order to bring about more flexibility in the analysis of data that do not fit the prerequisites pertaining to traditional taxonomies of research techniques. Recently, the convergence of ANOVA and MR has received considerable attention. (See, for example, Jennings (1967),Cohen (1968), Eikeland(1971a).) When a metric dependent variable can be regressed on both categoric and metric independent variables, a much more general multiple regression system has emerged.

The relationship between ANOVA and FA has also been discussed in the literature. Burt (1940) in his Factors of the Mind devoted a whole chapter to this problem. Most likely, he was the first to undertake such a comparison. Burt (1947), Creasy (1957), Bock (1960), and Gollob (1968) have also paid attention to overlappings of ANOVA and FA. All in all, these efforts do not seem to have resulted in bringing about a change in the way data are being analyzed, as judged by an almost complete lack of applications and discussions in the literature subsequent to these papers.

The reason why ANOVA is lagging behind as a factor analytic technique, and also as a regression system, seems to the author to be a result of the caprice of the historical development of ANOVA. ANOVA as a mathematical system and the sophisticated logic of experimental design were developed simultaneously. As a matter of fact, the same person, Ronald Fisher, is the father of both. This coincidence may have left the impression that ANOVA and experimental design are inseparable. ANOVA is to a very great extent conceived solely as a technique for making probabilistic statements about group differences. However, there is nothing wrong in separating ANOVA and experimental design. Rather, by seeing that the linkage is coincidental and not necessary, one is free to look for other uses of ANOVA in the analysis of data that are not obtained strictly the way the logic of experimental design prescribes.

ANOVA turns out to be a most powerful correlational technique. Measures of association can be developed, taking advantage of ANOVA as a machinery for assessing the possibility of complex relational systems in analyzing data involving multiple sources of variance. The potentiality of applying ANOVA in correlational analyses of complex data systems is indeed great and should be explored far more vigorously than hitherto.

Already Fisher himself realized that ANOVA could be used as a correlational technique. The correlation ratio (eta) goes back to his early work on ANOVA, as does the intraclass correlation. These two concepts, commonly associated with simple ANOVA designs, are basic to a further development of ANOVA as a more general correlational language.

The bivariate product-moment correlation technique has been used in the service of differential psychology for years. One must admit though that it has not served its function too well. The obvious drawback of this technique is that it can handle only two variables at a time. Differential psychology as concerned with explaining individual differences is in bad need for taking into account more than one source of variance at a time. For this purpose ANOVA is extremely well suited. It is so flexible that it can simultaneously decompose individual differences into group differences and inter- and intraindividual differences provided enough information is available.

This is far more than FA can do. So far developed, FA is mostly concerned with decomposing within groups differences into intraindividual sources of variance. That is: FA aims at partitioning the variance of test scores into linear combinations of observed variables or tests that can be interpreted as meaningful sources of variance. This same function can also be served by ANOVA. The difference between ANOVA and FA in this respect is that the former makes a priori linear combinations of tests, while the latter does this a posteriori.

Basic to an understanding of ANOVA as a correlational technique in differential psychology is to regard individual differences as constituting a distinct mode of classification. This means that individual differences have to be treated as a systematic source of variance. This is contrary to what is common practice in most experimental work where, as a rule, individual differences (within groups variance) are treated as error.

This report intends to make further explorations into ANOVA as a correlational language with a special view to the convergence on FA. Central to the discussion will be to show how the ANOVA concepts of eta (the correlation ratio) and alpha (the intraclass correlation) can be related to the two factor analytic models commonly met in the literature. They are variously named, like actual factors versus hypothetical factors (Nunnally 1967), observed versus inferred factors (Rozeboom 1966), or data reduction models versus classical factor analysis (Morrison 1967, Harman 1967). In the context of the present discussion a crucial distinction is made between *manifest* and *latent covariance structures*. It will be argued that the two classes of covariance structures bear a close relationship to the two factor analytic models.

Characteristic for ANOVA as an approach to factor analysis is that factors are defined prior to the analysis. The linear combinations of interest are fixed by the structure of the variables or tests. This a priori structure implies that the tests are grouped on a rational basis. Thus the analysis of such *test designs* can be seen as a hypothesis-testing procedure.

The conventional FA is based on a test design having the form of an N x k data matrix, where N denotes persons, or more generally observational units, and k denotes tests. In an N x k test design the k tests are undifferentiated, or not structured. If a set of k tests (or items) are grouped in order to measure one common trait, then the k tests are structured on a rational basis. A subsequent alpha analysis might well be called a rationalistic factor analysis.

An a priori structuring of tests could result in a test design like N x r x k. Here r denotes groups of assumedly homogeneous tests (or items), while k denotes, say, tests (or items) nested within each of the r groups of tests (or items).

Certainly, one could group observational units (persons) too, say, groups of persons somehow categorized. However, in order not to make the main features of the discussion too elaborate, no design will be imposed on persons in this report. (For an example of a worked-out multigroup-multifacet analysis, see Eikeland 1971b.)

2. ON THE CONVERGENCE OF ANALYSIS OF VARIANCE AND FACTOR ANALYSIS: SIMPLE CASE

2.1. Test design: N x 2

2.1.1. Eta analysis

The simplest data matrix that can be factor analyzed is an N x 2 matrix, say N persons and 2 tests. A full-rank solution means that one can form two orthogonal linear combinations of the two tests, the sum of the variances of the two linear combinations exhausting total variance. Total variance is here and in the following defined as the variance of the data matrix that is attributable to individual differences. In the case of an N x 2 matrix, total variance will be the sum of the variances of the sum scores and the difference scores.

For the purpose of the present exposition only the centroid solution among several factor analytic techniques is of concern. The reason why is that the centroid solution applies the same system of weights in forming linear combinations as does ANOVA, namely +1 and -1. By this system of weights there can be formed only one set of two orthogonal linear combinations for two tests. In contrast, by a principle components solution the system of weighting tests in linear combinations is completely liberalized, such that more than one set of two linear combinations, orthogonal to each other, is obtainable when two tests are being analyzed. By maximizing the variance of each linear combination extracted, the principle components analysis will have a unique solution. However, the weights are mathematically determined a posteriori, while the weights in the centroid solution are rationally based a priori. This distinction is important to keep in mind when a matching of ANOVA and FA is at issue.

a) Factor analyzing the variance-covariance matrix.

Let a 5 persons by 2 tests data matrix provide the data for subsequent analyses to be made. In TABLE 1 the 5 persons (P)

TABLE 1. 5 x 2 data matrix

P	T_1	T_2	Sum	Diff
a	5	4	9	+1
b	4	3	7	+1
c	3	5	8	-2
d	2	1	3	+1
e	1	2	3	-1

are identified by small letters from a to d, the two tests are denoted T_1 and T_2, the sum is each person's score across the two tests, and the difference (Diff) is each person's score on

TABLE 2. Variance-covariance matrix

	T_1	T_2
T_1	2,5	1,5
T_2	1,5	2,5

T_1 minus his score on T_2. The variance-covariance matrix of the data in TABLE 1 is presented in TABLE 2. Next, the two factors will be extracted according to the centroid method. The reader unfamiliar with this extraction procedure should consult a text-book in factor analysis, for example, Rozeboom (1966) or Nunnally (1967). In factor analyzing a variance-covariance matrix Rozeboom is particularly relevant. The factor matrix is presented

TABLE 3. Factor matrix of covariances

	F_1	F_2
T_1	$2^{\frac{1}{2}}$	$(0,5)^{\frac{1}{2}}$
T_2	$2^{\frac{1}{2}}$	$(0,5)^{\frac{1}{2}}$

in TABLE 3, where the covariances of the tests with the first and second normalized centroid factors are given. (Remember that a number raised to the power of $\frac{1}{2}$ equals the root of the same number.) By squaring and summing the entries in TABLE 3 for

TABLE 4. Factor variances and communalities

	F_1	F_2	h_i^2
T_1	$(2^{\frac{1}{2}})^2$	$((0,5)^{\frac{1}{2}})^2$	2,5
T_2	$(2^{\frac{1}{2}})^2$	$((0,5)^{\frac{1}{2}})^2$	2,5
V_{F_j}	4,0	1,0	5,0

rows and columns, communalities (h_i^2) and factor variances (V_{F_j}) are obtained. The total variance of the two tests, which is the sum of the principal diagonal in TABLE 2, also called the trace, is wholly explained by the two factors.

Applying the criterion of per cent total variance explained, or per cent trace, the first centroid factor accounts for 4/5 or 0,80; the second centroid factor accounting for 1/5 or 0,20.

b) Factor analyzing the correlation matrix.

The correlation between the two tests is given in TABLE 5. By factor analyzing the correlation matrix with 1's in the

TABLE 5. Correlation matrix

	T_1	T_2
T_1	1,0	0,6
T_2	0,6	1,0

principal diagonal, the factor matrix of TABLE 6 is obtained. The entries in TABLE 6 are the correlations between the tests

TABLE 6. Factor matrix

	F_1	F_2
T_1	$(0,8)^{\frac{1}{2}}$	$(0,2)^{\frac{1}{2}}$
T_2	$(0,8)^{\frac{1}{2}}$	$(0,2)^{\frac{1}{2}}$

and the factors. By squaring and summing columns one gets 1,6 and 0,4, which are the variances of the first and second factors, the two
or/linear combinations of the standardized data matrix. Again, by using the criterion of per cent trace, or per cent of total variance explained by the factors, the first centroid factor accounts for (1,6)/2, which equals 0,80. The second factor accounts for (0,4)/2, or 0,20. From TABLE 5 it can be seen that total variance, the trace, is 2. The result obtained by factoring the correlation matrix in terms of per cent trace for the two

factors is the same as the result obtained by factoring the variance-covariance matrix. It should be noted that this will not generally be the case. Only when the test variances are equal can the same results be obtained.

c) Variances of sum and difference scores.

In test theory the sum score for persons across tests is the most interesting linear combination. As a rule, it is the only combination of interest. This combination, as a matter of fact, is the same linear combination that is used in defining the first centroid factor. The variance of the sum scores in TABLE 1 is 8.

The other linear combination of the two test scores, the difference, is in test theory commonly used as a basis for defining error variance. This can be done only under the assumption that the two tests measure the same construct. Under this assumption the expected difference score should be zero. Therefore, an observed difference score is taken to mean random error of measurement. It was Rulon (1939) who first used the difference score for this purpose.

In FA, where a deterministic model is adopted, a difference score defines a linear combination of the tests and meets an acknowledged definition of a factor. In FA the difference score is interesting for the possibility of conveying information on individual differences.

When test variances are equal, the sum of raw scores and the difference of raw scores are uncorrelated. They thus define two orthogonal linear combinations, or factors. This is the case with the illustrating data in TABLE 1, where the sum of the variances of the two orthogonal linear combinations of test scores can be added to yield total variance. The variance of the difference scores in TABLE 1 is 2, and by adding the sum score variance of 8, the total variance equals 10.

The proportion of variance accounted for by the two linear combinations is 8/10, or 0,80, for the sum score and 2/10, or 0,20, for the difference score. It should be recalled that this is exactly the result obtained by the factor analytic procedure performed above.

d) Analysis of variance approach.

Next an ANOVA of the 5 x 2 data matrix in TABLE 1 will be performed in order to compare results obtained by this approach to results obtained by other methods above. The data matrix to

TABLE 7. ANOVA of 5 x 2 data matrix

Sources	SS	df	MS
Persons	16	4	4
Tests	0	1	0
P x T	4	4	1
Total	20	9	

be analyzed is in ANOVA terms a repeated measures design. The analysis is presented in TABLE 7 where SS is the sum of squares column, df degrees of freedom, and MS the mean squares for the three sources of variance. As there is no variation between the two tests across persons in the data, all variation is due to sources of variance connected with persons. Thus all variation is descriptive of individual differences. From TABLE 7 one can see that total sum of squares has been partitioned into an among persons sum of squares and an interaction sum of squares involving persons and tests.

In ANOVA the criterion score is partitioned into orthogonal effects. In the present case the score is explained by two effects that are linear combinations of the observed tests. The first one, an effect due to persons, is identified by summing the two tests. The second one, an effect due to a person by test interaction, is identified as a linear combination of the two tests formed by subtraction. That is, the interaction effect can be traced back to the difference score between the two tests. Here interaction and difference coalesce. This observation, however, is not easily made.

The data matrix in TABLE 1 can be decomposed into effect scores. This decomposition will be performed in order to show how the sums of squares in TABLE 7 are generated as the variation among partial criterion scores. The completely decomposed

TABLE 8. Decomposed data matrix

	T_1 (g p pt)	T_2 (g p pt)
a	3 + 1,5 + 0,5	3 + 1,5 - 0,5
b	3 + 0,5 + 0,5	3 + 0,5 - 0,5
c	3 + 1,0 - 1,0	3 + 1,0 + 1,0
d	3 - 1,5 + 0,5	3 - 1,5 - 0,5
e	3 - 1,5 - 0,5	3 - 1,5 + 0,5

data matrix is shown in TABLE 8. Here the observed scores are partitioned into three effect scores. The g effect is due to the grand mean. A common sense way of making this effect meaningful would be to say that the grand mean is the best prediction of each score in the set, provided no information is available such that persons and tests can be identified. The p effect is the difference between each person's average score across the two tests minus the grand mean. Lastly, the pt effect can here be regarded as the residual after having subtracted the g and the p effects from the observed score. The absolute values of the pt effects can also be seen as the average difference score.

It should be noted that since the g effect is a constant, it does not contribute to variance and is therefore ignored in the following. While the p effect naturally enough is equal for the two tests for each person, the pt effect has the same magnitude for the two tests for each person, but different signs. This can be interpreted to mean that the two tests measure something in common and something specific to each test. The specificity represents what is often called a bipolar factor. The bipolar characteristic is reflected in the fact that, say, person a scores relatively high on specificity in T_1; on the other hand, he scores relatively low on specificity in T_2.

From TABLE 8 one can compute directly the sums of squares attributable to each of the two effects by squaring and summing the partial scores, as they are all deviation scores. The two sums of squares add to total sum of squares of the data matrix. Thus all of the variation in the data matrix is taken care of by the p and the pt effects. The two sums of squares are 16 for p and 4 for pt, the total sum of squares being 20. This result is obtained by the ordinary procedure in TABLE 7 as well as by taking the sums of squares of the decomposed matrix in TABLE 8. By now taking the proportion of SS_p to SS_{total} one gets 16/20, or 0,80, which is the same per cent explained total variance for persons as found for the first centroid factor. Further, by taking the proportion of SS_{pt} to SS_{total} one gets 4/20, or 0,20, which is per cent total variance explained by the person by test interaction. This is the same amount of total variance as explained by the second centroid factor in the previous analysis. The sum of squares analysis performed here will be of particular interest as the discussion proceeds.

e) Concepts of variance of linear combinations.

In the N x 2 test design presently considered, the variance associated with persons is explained by two linear combinations, the sum and the difference of the two scores.

There are three conventional ways of computing the variance attributable to these sources. They yield different results, but are functionally related.

Consider the particular linear combination called the sum. In a psychometric tradition one usually computes the sum score variance,

$$V_X = (1/(N-1))\sum_{i=1}^{N}(x_{1i} + x_{2i})^2 = v_1 + v_2 + 2cov_{12}$$

where V_X is the sum score variance, N number of persons, x_{1i} and x_{2i} each person's deviation score for the two tests, v_1 and v_2 the variances for T_1 and T_2, and cov_{12} the covariance between the two tests.

In an ANOVA tradition the sum score variance is defined somewhat differently,

$$MS_p = (1/(N-1))\ 2\sum_{i=1}^{N}\left(\frac{x_{1i} + x_{2i}}{2}\right)^2 = (1/2)(v_1+v_2+2cov_{12})$$

where symbols are used as above. In addition, the multiplier of 2 and the denominator of 2 designate the two tests.

There is a third way to conceive the variance of the sum score, the variance of the average sum score,

$$V_{\bar{X}} = (1/(N-1))\sum_{i=1}^{N}\left(\frac{x_{1i} + x_{2i}}{2}\right)^2 = (1/4)(v_1+v_2+2cov_{12})$$

where no new symbols are introduced. It should be noted that the denominator in the multiplier of 1/4 in the derived formula is the number of tests squared.

In the context of the present discussion it is important to observe that MS_p in an ANOVA approach is 1/2 of the sum score variance, V_X, when two tests are used. The same relation will of course hold for the difference score variance, MS_{pt} is 1/2 of V_D.

Generally, the functional relationship among the three concepts of variance can be observed from the following equalities,

$$V_X = kMS_p = k^2V_{\bar{X}}$$

$$V_D = kMS_{pt} = k^2V_{\bar{D}}$$

Being familiar with these relationships should make the subsequent development of convergent features of ANOVA and FA more easily understandable.

f) Sums of squares and variances in a column-centered N x 2 matrix.

Taking proportions of sums of squares in order to find explained variance, as was done in section 2.1.1.d) above, seems somewhat strange, and ought to be explicated further. With a reference to the relationships established in the previous section, consider the ANOVA table of a column-centered

TABLE 9. ANOVA of a column-centered N x 2 data matrix.

Source	SS	df	MS	SS
Persons	SS_p	$(N-1)$	$MS_p = \bar{v}+cov$	$(\bar{v}+cov)(N-1) = SS_p$
P x T	SS_{pt}	$(N-1)$	$MS_{pt} = \bar{v}-cov$	$(\bar{v}-cov)(N-1) = SS_{pt}$
Total'	$SS_{t'}$	$2(N-1)$		$2\bar{v}(N-1) = SS_{t'}$

N x 2 data matrix as presented in TABLE 9. By a column-centered N x 2 matrix is here meant a matrix where the variation due to tests is partialled out. In doing this, 1 degree of freedom is used, and total degrees of freedom left is 2(N-1). In order to show that it is a partialled data matrix that is of concern, total sum of squares is denoted($SS_{t'}$).

It should be noted that $MS_p + MS_{pt} \neq MS_{t'}$. However, the two MS's can be added to yield the <u>total variance</u> of the two tests, or <u>trace</u>, the sum of the principal diagonal of the variance-covariance matrix.

It should also be noted that MS_p can be written as 1/2 of the variance-covariance matrix of the sum of the two tests, which is V_X: $MS_p = \frac{1}{2}V_X = \frac{1}{2}(2\bar{v} + 2cov) = \bar{v} + cov$.

It is not immediately obvious that $MS_{pt} = \bar{v}$ -cov. A proof of this identity was first given by Gulliksen (1950). The same proof is also given by Winer (1962).

From TABLE 9 it can be observed that if MS_p and MS_{pt} are each multiplied by (N-1), sums of squares are obtained, SS_p and SS_{pt}, respectively. Also, by adding MS_p and MS_{pt}, and multiplying the sum by (N-1) $SS_{t'}$, or the total sum of squares for the column-centered N x 2 data matrix, is obtained.

How much the variances of each of the two linear combinations of the two tests in TABLE 9 explain of the total variance can be found by taking the ratios of the MS's to total variance.

$$\frac{MS_p}{2\bar{v}} = \frac{MS_p(N-1)}{2\bar{v}(N-1)} = \frac{SS_p}{SS_{t'}} \quad \text{and} \quad \frac{MS_{pt}}{2\bar{v}} = \frac{MS_{pt}(N-1)}{2\bar{v}(N-1)} = \frac{SS_{pt}}{SS_{t'}}$$

In the language of FA, $2\bar{v}$ or $v_1 + v_2$ is a central concept, but not of much concern in an ANOVA context. It is the concept of trace, or the sum of the principal diagonal in the variance-covariance matrix, as mentioned before. Therefore, taking the proportions of the sums of squares of the two linear combinations of the two tests to the sum of squares of the column-centered data matrix, the corrected total sum of squares, amounts to finding per cent trace, which is an almost standardized procedure in FA.

g) <u>The sum of squares and sum of products matrices for an N x 2 test design.</u>

By a variance-covariance matrix in a test theoretic context is almost always meant a matrix for the sum of the tests . This must not necessarily be so. In the present discussion it will prove useful to extend the concept of a variance-covariance matrix to include matrices that are formed by any linear combination of the observed tests. For an N x 2 test design this means that one may have a variance-covariance matrix based on the difference scores as well. In TABLE 10 are presented

TABLE 10. SS & SP matrices for sum and difference scores

Sum

	x_1	x_2
x_1	10	6
x_2	6	10

Total=32

Difference

	x_1	$-x_2$
x_1	10	-6
$-x_2$	-6	10

Total=8

the sum of squares and the sum of products matrices (SS & SP matrices) for the two linear combinations of the two tests in the Nx 2 design, the sum and the difference. SS & SP matrices are here chosen in stead of variance-covariance matrices because a sums of squares analysis is at issue. The sum of the matrices in TABLE 10 is 32 and 8 for the sum and the difference, respectively. These are the sums of squares that one might obtain directly from the sum and difference scores in TABLE 1. By adding the two matrices the total sum of squares for the data matrix is obtained. This sum is 40, which compares to the SS_{tot}

for ANOVA in TABLE 7. According to the relationships found in 2.1.1.e) above the SS_{tot} in ANOVA should be ½ of 40, which is verified by TABLE 7.

By taking the proportions of the SS & SP matrices to the sum of the two matrices, 32/40 and 8/40, the same values obtain, 0,80 and 0,20, which are the per cent trace values obtained by the centroid FA procedure.

TABLE 11. SS & SP matrix for linear combinations

	(x_1+x_2)	(x_1-x_2)
(x_1+x_2)	32	0
(x_1-x_2)	0	8

Total = 40

Notice that by forming a SS & SP matrix of the two linear combinations, one obtains a diagonal matrix with a total sum of 40. That the matrix is diagonal obviously reflects the fact that the two linear combinations are orthogonal to each other. Again it should be recalled that orthogonality on the raw deviation score level is dependent on equal test variances. For convenience of illustrating purposes, homoscedasticity in hypothetical data has been deliberately sought.

h) Per cent trace and eta squared.

A squared correlation ratio, or an eta coefficient squared, is defined as the ratio of sum of squares among groups and total sum of squares in a simple ANOVA design. However, the concept of eta can be generalized to complex ANOVA designs, as shown by

Kennedy (1970) and Eikeland(1971a). The generality of the eta concept seems also to be corroborated by its applicability in an FA context, as closely related to more conventional measures within that tradition.

An eta coefficient can also be defined as the correlation between observed and predicted scores, the prediction being based on average values of score groups. In the test design presently considered, predicted scores are either an average sum or an average difference score. It is important to realize that these averages are linear combinations of the tests and are per definition factor scores. Therefore, the eta coefficient defined as the correlation between observed and predicted score, can also be considered a factor loading, i.e. the correlation between a test (observed score) and a linear combination of the tests (predicted score). By squaring and summing columns of the factor matrix, the variances of the linear combinations (factors) are obtained. We have shown for a very simple data matrix that the proportion of explained variance contributed by the linear combinations, per cent trace, is the ratio of sum of squares for groups to the total sum of squares for a column-centered matrix. These ratios are in effect eta squared coefficients. In the following much more will be said about eta as a relevant concept in FA.

i) ANOVA of a standardized N x 2 data matrix.

FA is usually performed on a standardized data matrix. This means that the tests are transformed to a scale with $\bar{X} = 0$ and $s = 1$. It also means that variances of linear combinations of the tests can be expressed in terms of 1's and r's, as these are the variances and covariances of the transformed tests.

ANOVA can also be applied to a standardized N x 2 data matrix. Before doing that, the variances of the two linear combinations of the two tests will be derived.

$$V_X = (1/(N-1))\sum_{i=1}^{N}(z_{1i} + z_{2i})^2 = (1/(N-1))(\Sigma z_{1i}^2 + \Sigma z_{2i}^2 + 2\Sigma z_{1i}z_{2i})$$

$$= 2 + 2r$$

$$V_D = (1/(N-1))\sum_{i=1}^{N}(z_{1i} - z_{2i})^2 = (1/(N-1))(\Sigma z_{1i}^2 + \Sigma z_{2i}^2 - 2\Sigma z_{1i}z_{2i})$$

$$= 2 - 2r$$

According to what was said about the relation between the variances of the sum and difference scores and the the corresponding MS's in ANOVA (see section 2.1.1.e)), it should be clear that the variances derived above have to be divided by 2 in order to obtain the MS's in an ANOVA table, which is presented

TABLE 12. ANOVA of standardized N x 2 data matrix

Source	SS	df	SS/(N-1)	SS
Persons	SS_p	N-1	$MS_p = 1 + r$	(1+r)(N-1)
P x T	SS_{pt}	N-1	$MS_{pt} = 1 - r$	(1-r)(N-1)
Total'	$SS_{t'}$	2(N-1)	2	2(N-1)

in TABLE 12. Again it should be noted that total variance, which in the context of the present discussion means trace, is 2, i.e. the sum of the principal diagonal in the correlation matrix. The $MS_{t'}$ certainly would be $(2(N-1))/(2(N-1)) = 1$, but this measure is of no concern here.

From TABLE 12 the proportions of variance explained by the linear combinations can be found. In the following the concept of eta squared will be substituted for per cent trace, or explained total variance.

$$eta_p^{\,2} = \frac{SS_p}{SS_{t'}} = \frac{(1+r)(N-1)}{2(N-1)} = \frac{1+r}{2}$$

$$eta_{pt}^{\,2} = \frac{SS_{pt}}{SS_{t'}} = \frac{(1-r)(N-1)}{2(N-1)} = \frac{1-r}{2}$$

j) <u>The N x 2 test design: Summary so far</u>.

Two linear combinations can be formed of two tests when the use of weights is restricted to +1 and -1. For the sum score the socalled design vector is +1 +1, for the difference score +1 -1. If the two tests have equal variances, the two linear combinations will be uncorrelated, or orthogonal to each other. By standardizing tests, homoscedasticity is imposed, and the sum and difference scores will be uncorrelated even when the tests in raw score form have unequal variances.

The convergence of the ANOVA concept of eta squared on the FA concept of per cent trace has been explicated. For this particular test design it has been shown that FA (as a data reduction model) corresponds to ANOVA on the level of a sums of squares analysis.

2.1.2. Alpha analysis

So far the concern has been to form linear combinations of observed test scores and to partition total variance into observed variance-covariance matrices based on these linear combinations (se Section 2.1.1.g)). This is characteristic for ANOVA on the sums of squares level as well as for FA when actual factors, or straightforward linear combinations of tests, are involved, and not hypothetical factors. The FA method of concern in the preceding sections has been the centroid solution with variances in the principal diagonal of the variance-covariance matrix, or 1,s in the principal diagonal of the correlation matrix.

The FA model as used in a more strict sense, however, is more concerned with inferred variables, rather than observed variables. This means among other things that one is trying to explain observed covariances among variables by making inferences to underlying common traits. This is the Spearman-Thurstone tradition in FA. The analysis starts with making estimates of communalities in the principal diagonal of the variance-covariance matrix, or the correlation matrix, in order to try to separate what the variables (tests) share with the other variables (tests) from what is unique to each variable (test).

While the analysis discussed in Section 2.1.1. may be said to be concerned with manifest variance-covariance matrices, the present section will be concerned with developing the concept of a latent variance-covariance matrix, which seems to underly the theory of making inferences both in ANOVA and in FA. It will be argued that alpha as an intraclass correlation coefficient[a] can be used as a general measure in exploring the latent covariance structures to be dealt with in the following. The intention is also to contrast alpha as a generic concept for latent covariance structures to eta as a generic concept for manifest covariance structures.

a) A variance components analysis of a 2 x 2 variance-covariance matrix.

There seems to be an ANOVA parallel to FA in the Spearman-Thurstone sense. In what is called a variance components analysis, one is concerned with partitioning observed variances into inferred sources of variance.

Again it may prove fruitful to bring in Rulon's (1939) partition of total test variance into sum and difference score variance. Rulon defined difference score variance as random variance, or error of measurement variance, as conceived within the theory of homogeneous tests. In classical test theory one has been almost exclusively concerned with one-factor composites. In that context Rulon imposed the random component variance on the observed sum score variance to separate what might be considered signal and what noise in this source of variance. As is well known, in test theory the observed score variance is conceptually composed of a true score component and an error component.

A variance components analysis of an N x 2 data matrix is performed in TABLE 13, where some modifications are made compared to what is conventionally done in writing out an ANOVA table. The table is identical to TABLE 9 as far as the first

TABLE 13. Variance components analysis of an N x 2 matrix

Source	SS	df	MS	Model	VC
Persons	SS_p	(N-1)	MS_p	$\sigma^2_{pt} + 2\sigma^2_p$	σ^2_p
P x T	SS_{pt}	(N-1)	MS_{pt}	σ^2_{pt}	σ^2_{pt}
Total'	$SS_{t'}$	2(N-1)			$\sigma^2_{pt} + \sigma^2_p$

four columns are concerned. While the MS's in TABLE 9 were identified with observed variance-covariance matrices, the MS's in TABLE 13 have been modelled according to an inferred variance structure. In an ordinary ANOVA context, where probabilistic statements are sought, the table would have an E(MS) column. Instead of the commonly used E(MS) column, TABLE 13 has a column called Model to emphasize that the analysis is being performed for a descriptive purpose. Thus the models, as will be shown shortly, can be regarded as theoretically structured variance-covariance matrices in terms of variance components. The MS's of TABLE 9 describe manifest variance-covariance matrices. The MS's of TABLE 13 describe latent variance-covariance matrices, in that they are not observed, but inferred.

The VC column of TABLE 13 designates the unweighted variance components. The sum of the two unweighted components, $\sigma^2_{pt} + \sigma^2_p$, is the inferred structure of one average test's variance.

What the models in TABLE 13 imply, can be made quite concrete by presenting the variance-covariance matrices based on the theoretical structures of the variances for the two linear combinations of the two tests, the sum score and the difference score. The latent variance-covariance matrix for the sum is pre-.

TABLE 14. Latent variance-covariance matrix of sum

	x_1	$+x_2$
x_1	$\sigma^2_{pt} + \sigma^2_p$	σ^2_p
$+x_2$	σ^2_p	$\sigma^2_{pt} + \sigma^2_p$

Total $= 2\sigma^2_{pt} + 4\sigma^2_p = V_X$

TABLE 15. Latent variance-covariance matrix of difference

	x_1	$-x_2$
x_1	$\sigma^2_{pt} + \sigma^2_p$	$-\sigma^2_p$
$-x_2$	$-\sigma^2_p$	$\sigma^2_{pt} + \sigma^2_p$

Total $= 2\sigma^2_{pt} = V_D$

sented in TABLE 14. The corresponding matrix for the difference is presented in TABLE 15.

In order to see the reltionship[a/] between the two latent variance-covariance matrices of TABLES 14 and 15 and the models for the MS's in TABLE 13, it should be recalled what was found in Section 2.1.1.e concerning the functional relationships between the different ways of computing variances of linear combinations. There it was shown that the sum of the variance-covariance matrix, i.e. the variance of the sum score, is k times MS_p. Thus by multiplying the model for MS_p in TABLE 13 by 2, the following result is obtained

$$V_X = 2MS_p = 2(\sigma^2_{pt} + 2\sigma^2_p) = 2\sigma^2_{pt} + 4\sigma^2_p$$

which is the total of the latent variance-covariance matrix of TABLE 14. The same will of course hold for MS_{pt} and the total of the latent variance-covariance matrix for the difference score.

What is most interesting about TABLE 14 (and TABLE 15) is the partitioning of the test variance into a covariance component, σ_p^2, and a unique component, σ_{pt}^2. In this particular design the unique component is conceptually a confounding of specificity and error of measurement as the design is an unreplicated one: There is only one measure within each person-by-test cell. The decomposition of the test variance thus satisfy the requirement of FA proper that the variance should be conceived as being composed of a common source and a unique source. (For a comparison between the FA and the classical test theory decomposition of variance, see Lord & Novick 1968, Chapter 24.)

By now the parallel to FA should be relatively easy to grasp. From TABLE 14 one should be able to see that in order to find the proportion of variance that is explained by the covariance component, he can take the ratio

$$\frac{\sigma_p^2}{\sigma_{pt}^2 + \sigma_p^2} = \frac{2\sigma_p^2}{2\sigma_{pt}^2 + 2\sigma_p^2} = \frac{MS_p - MS_{pt}}{MS_p + MS_{pt}} = \rho_I = \text{alpha}_{(1)}$$

which in effect is a per cent trace measure. This means that for the average test one asks how much of the variance is explained by the covariance, the component, or the factor. Certainly, one could do the same for the interaction component to see how much that component explains. (However, this is not too interesting from a substantive point of view as the interaction component is a confounding of specificity and error.) From the formula above one can see that FA is being linked to the concept of the intraclass correlation, ρ_I, as the first part of it would be the defining formula of that concept. It should also be recognized that what is here taken as a per cent trace measure is also the definition of the reliability of one average test, alpha$_{(1)}$.

b) FA and the Spearman-Brown rationale.

In TABLE 14 there is even more information on explained variance than is commonly extracted by an ordinary FA procedure. If one asks how much of the sum score variance, instead of how much of the average score variance, is explained by the first centroid factor, he obviously should be interested in the ratio

$$\frac{4\sigma_p^2}{2\sigma_{pt}^2 + 4\sigma_p^2} = \frac{2\sigma_p^2}{\sigma_{pt}^2 + 2\sigma_p^2} = \frac{MS_p - MS_{pt}}{MS_p} = 1 - \frac{MS_{pt}}{MS_p} = \text{alpha}_{(2)}$$

While the formula in Section 2.1.2.a (the preceding section) is concerned with explained variance in one test, or more correctly, one average test, the present formula is concerned with a possibly even more interesting ratio, the proportion of explained sum score variance, since the sum score very often is the actual measure used in making decision about people. The above formula will be recognized by the informed reader as the well-known Hoyt-Cronbach coefficient alpha.

In the present context it should be extremely useful to translate the traditional alpha form as an estimate of reliability into an FA language: Coefficient alpha is a measure of how much the first centroid factor explains of the sum score variance.

It is important to be aware of the fact that FA, as far as the author knows, never paid attention to this aspect of an N x 2 or, more generally, an N x k matrix. It seems that FA has not been able to include in its theory the Spearman-Brown rationale, i.e. how to account for sum score variance. Neither in the very simple case of an N x 2 matrix, nor in the case of an N x k matrix is there a traditional machinery available to tease out this information.

c) Variance components analysis of a 2 x 2 correlation matrix.

According to what is said so far concerning an N x 2 matrix, there are three different ways of considering an FA of a variance-covariance matrix. The three modes of analysis can be

TABLE 16. Variance components analysis of an N x 2 standardized matrix.

Source	SS	df	MS	VC
Persons	SS_p	(N-1)	MS_p 1 + r	r
P x T	SS_{pt}	(N-1)	MS_{pt}= 1 - r	1 - r
Sum	$SS_{t'}$	2(N-1)	2	1

summarized by examining a correlation matrix instead of a variance-covariance matrix of raw scores. The parallel between analyzing a correlation matrix and a variance-covariance matrix should be understood. In TABLE 16 the variance components analysis of a standardized N x 2 matrix is presented. In that table one has gone one step further compared to TABLE 12, in that the components have been derived in terms of functions of the correlation matrix. This has been accomplished by drawing on the structural model for this particular test design as presented in TABLE 13. It should be noted that in TABLE 16 the components, σ_p^2 and σ_{pt}^2, are written as r and 1-r, respectively, in order to keep close to the conventional language for a correlation matrix.

As emphasized above, a variance components analysis elaborates on an inferred variance structure in that one goes beyond the observed variances of linear combinations by imposing a theoretical structure on the observations. The inferred variance structure can be utilized in constructing a latent correlation matrix for each of the two linear combinations, like what has been done in TABLES 14 and 15 with the variance-covariance matrices for the same linear combinations. The latent correlation

TABLE 17. Latent correlation matrices for sum and difference.

	z_1	$+z_2$
z_1	(1-r)+r	+r
$+z_2$	+r	(1-r)+r

	z_1	$-z_2$
z_1	(1-r)+r	-r
$-z_2$	-r	(1-r)+r

matrices are shown in TABLE 17. It should be noted that the sums of the matrices are unchanged: The variance for the sum is 2 + 2r, the variance for the difference 2 - 2r. By dividing each of these variances by 2, the number of tests, one gets 1+r and 1-r, which are the MS's obtained in TABLE 12 and TABLE 16. The total variance is the sum of the two matrices in TABLE 17, (2 + 2r) + (2 - 2r) = 4.

Consider now a first set of ratios, illustrated by the data presented in TABLE 1:

$$\frac{2+2r}{4} = \frac{1+r}{2} = \frac{1,6}{2} = 0,80 \qquad (eta_p{}^2)$$

$$\frac{2-2r}{4} = \frac{1-r}{2} = \frac{0,4}{2} = 0,20 \qquad (eta_{pt}{}^2)$$

Here is first shown the ratio of observed sum score variance to total variance, or trace. This analysis may be regarded as taking the ratio of SS_p to $SS_{tot'}$, yielding an eta squared coefficient, as shown in the discussion following TABLE 12. The second ratio involves the difference score variance: It equals the ratio of SS_{pt} to $SS_{tot'}$. Also, the two ratios are squared product-moment correlations among observed scores and predicted scores, the predictions being based on two orthogonal linear combinations of the observed scores.

Consider next another set of ratios, based on unweighted variance components:

$$\frac{r}{(1-r)+r} = \frac{r}{1} = r = 0,60 \qquad (alpha_{p(1)})$$

$$\frac{1-r}{(1-r)+r} = \frac{1-r}{1} = 0,40 \qquad (alpha_{pt(1)})$$

Here are shown the ratios of inferred common and unique variances to the observed variance of one test. As is evident from TABLE 17 the one-test variance of unity has been decomposed into a common component and a unique one. In test theory the two ratios are the reliabilities of one (average) test and the difference score between the two tests. In the case of the difference score reliability, the measure is not too meaningful for the present test design, which is an unreplicated one such that specificity and error are confounded.

Consider a third set of ratios formed from TABLE 17:

$$\frac{4r}{2 + 2r} = \frac{2r}{1 + r} = \frac{1,2}{1,6} = 0,75 \qquad (\text{alpha}_{p(2)})$$

$$\frac{2(1 - r)}{2 + 2r} = \frac{1 - r}{1 + r} = \frac{0,4}{1,6} = 0,25 \qquad (\text{alpha}_{pt(2)})$$

The first coefficient is the proportion of the total correlation matrix, i.e., the sum score variance of standardized tests, that is inferred common variance. This coefficient is the reliability of the sum score. 75 % of sum score variance is the estimated proportion of variance accounted for by the common component. The unique component accounts for 25 % of sum score variance.

Notice that the first formula above is the Spearman-Brown prophecy formula (for double length), derived from the latent structure of the correlation matrix.

2.2. Test design: N x k

Next, we are going to see to what extent the results obtained so far for an N x 2 test design can be generalized to an N x k design.

2.2.1. Eta analysis

a) The variance-covariance matrix of an N x k test design.

Consider a column-centered data matrix with N rows and k columns, where N designates persons and k tests or variables.

In a column-centered N x k data matrix the variance among tests has been partialled out. Therefore, the total test variance that is left is the sum of the test variances, i.e., Σv_i or $k\bar{v}$. What

TABLE 18. ANOVA of an N x k data matrix.

Source	SS	df	SS/(N-1)	Structural model
Persons	SS_p	$(N-1)$	$MS_p = \bar{v}+(k-1)\overline{cov} =$	$\sigma^2_{pt} + k\sigma^2_p$
P x T	SS_{pt}	$(N-1)(k-1)$	$(k-1)MS_{pt} = (k-1)(\bar{v}-\overline{cov}) =$	$(k-1)\sigma^2_{pt}$
Sum	$SS_{t'}$	$k(N-1)$	$k\bar{v}$	$k\sigma^2_{pt} + k\sigma^2_p$

is evident from TABLE 18 is that there are (k-1) linear combinations in the interaction term. These linear combinations can be regarded as (k-1) differences or contrasts. A set of (k-1) orthogonal contrasts can be found among many others that will completely account for the variance attributable to the PT source of variance.

In the present test design a full-rank solution will mean that one can extract k factors or linear combinations, say, one sum (the P source of variation) and (k-1) contrasts (the PT source of variation). The total variance in the column-centered N x k matrix can thus be transformed to the sum of 1 + (k-1) variances of orthogonal linear combinations of the k original tests.

In ANOVA the (k-1) contrasts are conventionally lumped together. In FA one would be interested in the separate contrasts for possible interesting information. Certainly, this can also be accomplished in ANOVA, if need be.

From TABLE 18 can be found how much of total variance is accounted for by the sum score, or person variance; and how much by the (k-1) contrasts, or the person by test interaction. In terms of variances and covariances, the following ratios are informative,

$$eta_p^2 = \frac{\bar{v} + (k-1)\overline{cov}}{k\bar{v}}$$

$$eta_{pt}^2 = \frac{(k-1)(\bar{v} - \overline{cov})}{k\bar{v}}$$

The ratios above are obviously per cent trace measures since $k\bar{v}$ equals trace. It should also be clear that the ratios are eta squared coefficients, as they can be written,

$$eta_p^2 = \frac{MS_p(N-1)}{k\bar{v}(N-1)} = \frac{SS_p}{SS_t},$$

$$eta_{pt}^2 = \frac{MS_{pt}(k-1)(N-1)}{k\bar{v}(N-1)} = \frac{SS_{pt}}{SS_t},$$

b) Analyzing a k x k correlation matrix.

Consider the k x k correlation matrix of an N x k test design. The parallel to analyzing the variance-covariance matrix should be relatively easy to see.

TABLE 19. ANOVA of standardized N x k matrix.

Source	SS	df	SS/(N-1)
Persons	SS_p	$(N-1)$	$MS_p = 1 + (k-1)\bar{r}$
P x T	SS_{pt}	$(N-1)(k-1)$	$(k-1)MS_{pt} = (k-1)(1-\bar{r})$
Sum	SS	$k(N-1)$	k

Analyzing the k x k correlation matrix can be done by performing an ANOVA of the standardized N x k data matrix. This means that each of the test scales is transformed to a standard score scale. By this transformation the data matrix is being column-centered, and there will be no variance among tests.

The variances of the linear combinations can be found in TABLE 19. Again we have partitioned the sum of the test variances (total variance) into variances of k linear combinations, one sum and (k-1) contrasts or differences. In order not to complicate matters too much at this moment, the (k-1) contrasts are lumped together. Later it will be shown how they can be separated.

The two proportions of variance accounted for, or the per cent trace coefficients are now,

$$eta_p^2 = \frac{1 + (k-1)\bar{r}}{k}$$

$$eta_{pt}^2 = \frac{(k-1)(1-\bar{r})}{k}$$

From TABLE 19 it should be pretty clear that the ratios above are eta squared coefficients, as they can be shown to be ratios of sums of squares.

2.2.2. N x k test design: Relation of ANOVA to FA.

The eta_p^2 above, i.e., the proportion of variance accounted for by the sum score, will be examined somewhat closer in order to explore the relation to FA, centroid solution.

Let R be the sum of the correlation matrix; and Σr_{ig} the sum of the correlations of each of the tests with the sum of all k tests, i.e., the sum of the columns in the correlation matrix. The development of the following relationship between ANOVA and FA seems to be of a fundamental character:

$$\text{eta}_p^{\,2} = \frac{1 + (k-1)\bar{r}}{k} = \frac{k + k(k-1)\bar{r}}{k^2} = \frac{R}{k^2} = \left(\frac{R^{\frac{1}{2}}}{k}\right)^2 = \left(\frac{R}{kR^{\frac{1}{2}}}\right)^2$$

$$= \left(\frac{\Sigma r_{ij}}{kR^{\frac{1}{2}}}\right)^2, \text{where } i=j,\ i \neq j$$

$$= \left((1/k)\frac{\Sigma r_{ij}}{R^{\frac{1}{2}}}\right)^2 = \left(\frac{\Sigma r_{ig}}{k}\right)^2 = (\overline{r_{ig}})^2$$

The derivation shows that eta squared for persons equals the square of the average "test/test sum" correlation. As r_{ig} denotes loadings, we have shown that the average loading squared is identical to eta squared for persons.

It is of crucial importance to notice the subtle discrepancy between ANOVA and FA here. In extracting the first centroid in FA, one first finds the correlation between each of the tests with the sum of the tests, i.e., the factor loadings. Then one sums the squares of the loadings and averages. Thus, in FA one is concerned with $(\Sigma r_{ig}^{\,2})/k = \overline{r_{ig}^{\,2}}$, while in ANOVA one ends up with $(\overline{r_{ig}})^2$. The relation between ANOVA and FA as formally shown here was pointed to by Burt (1940), Chapter 10. It is indeed interesting to notice the different ways of turning out the result. One is left wondering whether the one way is more appropriate than the other.

If all correlations in a matrix are equal, then certainly $\overline{r_{ig}^2} = (\overline{r_{ig}})^2$. However, if the correlation coefficients are not equal, and this is the realistic case, then $\overline{r_{ig}^2} > (\overline{r_{ig}})^2$. Generally, the following relation holds, $\overline{r_{ig}^2} \geq (\overline{r_{ig}})^2$.

Thus, the proportion of variance accounted for by the first centroid factor as extracted by an FA procedure is either equal to or greater than the proportion of variance accounted for by an ANOVA procedure. The same relation should not be generally valid for extraction of subsequent factors, as there is a fixed amount of variance to be accounted for.

2.2.3. Alpha analysis.

a) From the covariance matrix.

In TABLE 18 MS_p and $(k-1)MS_{pt}$ are also given in terms of variance components, which may be regarded as elements of a latent covariance matrix (see TABLE 14). The same ratios as developed in 2.1.2.a) for the N x 2 test design can be obtained for the N x k design. In addition to the eta squared coefficients given in 2.2.1.a), two sets of alpha coefficients can be found. The first set concerns the proportion of common and unique contribution to one average test variance. This way of structuring the test variance should be a genuine parallel to the Spearman-Thurstone tradition in FA. It should be noted that the latent variance-covariance matrix for the N x k test design will be different from the corresponding variance-covariance matrix for the N x 2 test design as shown in TABLE 14, in that there will be k(k-1) off-diagonal cells in the extended design compared to 2 for the simplest one.

However, the structure of the principal diagonal will be exactly like for the two matrices. The reason why is that no additional component or factor is assumed for the extended design. The assumption is still that the k tests are in some sense homogeneous; that is, they are meant to measure one underlying trait.

Therefore, as the average test variance for the N x k test design is conceived to have the structure, $\sigma^2_{pt} + \sigma^2_p$, i.e., being composed of a unique and a common component, the proportion each of them contributes to the average test variance is easily obtained, like what has been done in 2.1.2.a).

On the other hand, when the contribution to the sum score variance made by the unique and common components is involved, one has to take into account that the contribution the common component increases much faster than the contribution of the unique component. As a matter of fact, the common component will increase by a factor of k^2, while the unique component will increase by a factor of k. This follows from the theoretical structure of the latent variance-covariance matrix for an N x k test design, which has the form,

$$V_X = k\sigma^2_{pt} + k^2\sigma^2_p$$

where V_X, the sum of the matrix, is the observed variance of the sum score across all of the k tests. As shown in 2.1.1.e), MS_p is 1/k of V_X, i.e., $\sigma^2_{pt} + k\sigma^2_p$, which is the structural model given in TABLE 18.

In giving the proportions of unique and common components to the sum score variance, one has to use weighted components instead of unweighted ones. The proportions obtained are in form alpha coefficients, as they may be regarded as latent trait measures.

$$\text{alpha}_{p(k)} = \frac{k\sigma_p^2}{\sigma_{pt}^2 + k\sigma_p^2} = \frac{MS_p - MS_{pt}}{MS_p}$$

$$\text{alpha}_{pt(k)} = \frac{\sigma_{pt}^2}{\sigma_{pt}^2 + k\sigma_p^2} = \frac{MS_{pt}}{MS_p}$$

Here $\text{alpha}_{p(k)}$ gives the proportion of sum score variance that is due to an inferred common trait across the k tests, while $\text{alpha}_{pt(k)}$ indicates how much of the sum score variance is influenced by a unique contribution of tests.

The informed reader will of course recognize $\text{alpha}_{p(k)}$ as traditional alpha in a test theoretical context. However, the intention here is to show how close the generic concept of alpha is to a factor analytic consideration.

It should be understood that the present mode of analyzing the variance-covariance matrix in terms of weighted components has no parallel in FA. As previously remarked, FA has never been able to link its rationale to the Spearman-Brown rationale in test theory. The step from an unweighted components analysis to a weighted components analysis seems to indicate how this gap can be bridged. There is more information in the variance-covariance matrix than FA so far has managed to exploit.

b) From the correlation matrix.

From TABLE 19 and the structural model in TABLE 18 one can go a step further to find the components in terms of $\bar{r}$, the average intertest correlation. By so doing, one obtains $\sigma_p^2 = \bar{r}$, and $\sigma_{pt}^2 = 1-\bar{r}$. These are the same results as obtained for the N x 2 test design, except for the average correlation coefficient that is needed in the extended design where altogether $\frac{1}{2}k(k-1)$ different correlation coefficients among tests exist, compared to the single one in the simplest design.

It should be recalled that the sum of the unweighted components for standardized scores is unity; i.e., $\sigma_{pt}^2 + \sigma_p^2 = (1-\bar{r}) + \bar{r} = 1$. Thus, $\bar{r}$ indicates directly the estimated proportion of common variance in one average test, while uniqueness is indicated by $1-\bar{r}$.

The structural model for MS_p, $\sigma_{pt}^2 + k\sigma_p^2$, can be written as a function of the correlation matrix when standardized scores are being analyzed: $MS_p = (1-\bar{r}) + k\bar{r}$. This is the variance of the sum of k standardized scores, and the proportions of common and unique variance can be given as alpha coefficients:

$$\text{alpha}_{p(k)} = \frac{k\bar{r}}{(1-\bar{r}) + k\bar{r}} = \frac{k\bar{r}}{1 + (k-1)\bar{r}}$$

$$\text{alpha}_{pt(k)} = \frac{(1-\bar{r})}{(1-\bar{r}) + k\bar{r}} = \frac{(1-\bar{r})}{1 + (k-1)\bar{r}}$$

It is interesting to notice that by the procedure adopted here $\text{alpha}_{p(k)}$ emerges in the form of the general Spearman-Brown prophecy formula.

2.2.4. N x k test design: Numerical example.

We are going to show the different ways of "factor" analyzing a data matrix by way of a 5 x 4 matrix. A small matrix is chosen in order to keep data as simple as possible. The data matrix and the variance-covariance matrix are presented in TABLES 20 and 21, where the 5 persons are symbolized with small letters from a - e, and the 4 tests with numerals from 1 - 4.

TABLE 20. A 5 x 4 data matrix.

	1	2	3	4
a	5	5	4	4
b	2	3	5	4
c	4	3	3	2
d	2	3	1	2
e	1	2	3	2

TABLE 21. Variance-covariance matrix of the 5 x 4 data matrix.

	1	2	3	4
1	2,70	1,55	0,55	0,70
2	1,55	1,20	0,45	0,80
3	0,55	0,45	2,20	1,30
4	0,70	0,80	1,30	1,20
Sum	5,50	4,00	4,50	4,00

a) Extracting the first centroid.

In order to be able to compare results with an ordinary ANOVA, the variance-covariance matrix will be analyzed and not the correlation matrix.

The loadings in terms of covariances are found by the traditional procedure,

	F_1
T_1	$5,5/(18^{\frac{1}{2}})$
T_2	$4,0/(18^{\frac{1}{2}})$
T_3	$4,5/(18^{\frac{1}{2}})$
T_4	$4,0/(18^{\frac{1}{2}})$

When these loadings are squared and summed, one gets 4,58, which is the variance of the first centroid factor.

One could continue extracting the remaining three factors in order to have a full-rank solution. However, instead of doing this we are going to lump the variances of the three factors into one residual variance. This is a way of simplifying the analysis. Nothing is lost for our illustrating purpose.

The total variance of the column-centered data matrix is 7,3. This is the trace of the variance-covariance matrix of TABLE 21. With a variance of 4,58 due to the first centroid factor, there remains 7,30 - 4,58 = 2,72 as residual variance.

In terms of per cent trace, the following results are obtained:

$$F_1 = 4,58/7,3 = 0,627$$
$$F_{res} = 2,72/7,3 = 0,373$$

These results will be compared to the subsequent analyses to be performed.

b) Eta analysis (Sums of squares analysis).

An ordinary ANOVA is presented in TABLE 22. It should be remembered that total sums of squares marked means that the variance due to tests is ignored. With a column-centered data matrix the concern is only the variance attributable to individual differences. Remember also that the person by test inter-

TABLE 22. ANOVA of the 5 x 4 data matrix.

Source	SS	df	SS/(N-1)	
Persons	18,0	4	4,5	(MS_p)
P x T	11,2	12	2,8	($(k-1)MS_{pt}$)
Total SS =	29,2		7,3	(trace)

action is given as a sum of a set of three orthogonal linear combinations of the tests, i.e., $SS_{pt}/(N-1) = (k-1)MS_{pt}$. By this procedure total test variance can be accounted for.

From TABLE 22 per cent trace measures can be found. They are also eta squared coefficients:

$$Eta_p^2 = 4,5/7,3 = 18,0/29,2 = 0,616$$
$$Eta_{pt}^2 = 2,8/7,3 = 11,2/29,2 = 0,384$$

As can be seen, the results obtained by the two factor analyzing procedures are not exactly equal. The discrepancy observed, however small, was expected according to the relations revealed in 2.2.2.

c) <u>Alpha analysis (Variance components analysis)</u>.

A variance components analysis implies that one is considering inferred variables rather than observed ones. A theoretical structure is imposed on the observed test variance. Thus, according to the structural model as presented in TABLE 18, the average test variance of 1,825 in TABLE 23 is composed of a

TABLE 23. Variance components analysis of 5 x 4 data matrix.

Source	SS	df	MS	VC	Structural model
Persons	18,0	4	4,5	0,892	0,933 + 4 . 0,892
P x T	11,2	12	0,933	0,933	0,933
	29,2			1,825	

common component (common to all 4 tests) and an average test-unique component: 1,825 = 0,892 + 0,933. From this one can find how much of the average one-test variance is explained by the common component and how much by the unique component:

$$\text{alpha}_{p(1)} = 0,892/1,825 = 0,489$$
$$\text{alpha}_{pt(1)} = 0,933/1,825 = 0,511$$

The above results are based on an unweighted components analysis. A weighted components analysis can be performed for the sum score variance, or the person variance. From TABLE 23 one can see that according to the structural model the observed MS_p is composed of a unique component with weight 1 and a common component with weight 4. Thus, the structural model gives the clue to finding how much of this observed variance is estimated due to the common factor and how much due to the unique factor:

$$\text{alpha}_{p(4)} = 3,568/4,5 = 0,793$$
$$\text{alpha}_{pt(4)} = 0,933/4,5 = 0,207$$

It is certainly of considerable interest to learn how much of the sum score variance is explained by the first factor, the common factor. This is what has been of concern in test theory, where the one-factor test has constituted the basis for theorizing. As pointed to above, traditional FA never reached this level of analysis.

d) Forming four linear combinations of the four tests.

According to the theory of the single-factor test composite the residual in the present example would be considered error variance. Although the residual could be partitioned into three factors or linear combinations, each of them according to theory should convey noise and not signal.

One certainly frequently questions that test composites behave according to theory. Factor theory has a way of examining the residual in order to see whether more signals are present. ANOVA can also do this.

The residual sum of squares of 11,2 in the present example can be partitioned into three orthogonal components in addition to the one formed by the sum score. Consider the following design matrix of signs and weights as a point of departure in forming four linear combinations of the four tests:

	T_1	T_2	T_3	T_4
I	+1	+1	+1	+1
II	+1	+1	−1	−1
III	+1	−1	+1	−1
IV	+1	−1	−1	+1

The linear combinations are denoted by I, II, III, IV. These four linear combinations meet the requirement for forming orthogonal contrasts. By applying the signs and weights above to the data matrix of TABLE 20, the "factor" scores for each of the 5 persons can be found, and the variances of the linear combinations computed. This has been done in TABLE 24. The variances

TABLE 24. Variances of the four orthogonal linear combinations.

	I	II	III	IV	
a	18	2	0	0	
b	14	-4	0	-2	
c	12	2	2	0	
d	18	2	-2	0	
e	8	-2	0	-2	
V(X)	18,0	8,0	2,0	1,2	Sum =29,2
MS(X)	4,5	2,0	0,5	0,3	Sum = 7,3
			2,8 (II–IV)		

of the four linear combinations add to total variance of the four tests, a consequence of the orthogonality established among the combinations. It should be noted that the residual variance of 2,8 (see TABLE 22) has now been partitioned into three variances.

The analysis performed in TABLE 24 is an example of a full-rank solution in an ANOVA context. What has been done here, is to find just one set of many possible sets of four linear combinations, orthogonal to each oether, that could be formed. The criterion for which set of linear combinations to choose as most interesting in a substantive sense, is different for FA and the present approach, ANOVA.

One very important distinction between FA and ANOVA is that FA makes a posteriori linear combinations, often based on mathematical criteria; while ANOVA is based on a priori linear combinations, following rational considerations.

Complex test designs often implies such rational linear combinations. The design matrix above is typical of groups of tests that might be combined on a multifacet basis. One can conceive the four tests as forming a system with two modes of classifications with two levels for each of them, such that a 2 x 2 factorial test design emerges. The first linear combination explores the common aspect of the two modes of classification, the second explores the contrast between the levels of the first mode of classification, the third the contrast between the levels of the second mode of classification. Lastly, the fourth linear combination explores the interaction of the two modes of classification. These linear combinations are substantively meaningful prior to the analysis. The whole design is constructed with a view to testing just these linear combinations. Thus, factor analysis by ANOVA will be a rationalistic analysis with a much more theoretical slant than most of traditional FA.

A closer examination of the character of an a priori factor analysis by way of an ANOVA procedure will be undertaken in Part 3.

3. FACET FACTOR ANALYSIS: COMPLEX CASE[1]

3.1. N x 2 x 2 factorial test design.

Consider the following design of a battery of four tests: There are two verbal and two performance tests, say. The verbal tests are an opposites and a similarities test. The same is the case for the performance tests. These four tests form what may

TABLE 25. Data matrix of a 2 x 2 factorial test design.

	11	12	21	22
1	X_{111}	X_{112}	X_{121}	X_{122}
2	X_{211}	X_{212}	X_{221}	X_{222}
.	.	.	.	.
.	.	.	.	.
n	X_{n11}	X_{n12}	X_{n21}	X_{n22}

called a 2 x 2 factorial test design. The design is shown in TABLE 25, where the column headings symbolize the categories of tests: 11 is the verbal-opposites test, 12 the verbal-similarities, 21 the performance-opposites, and 22 the performance- similarities test. The index n is the number of persons tested.

The test design is factorial or crossed, in that all possible combinations of the two traits (verbal and performance) with the two formats (opposites and similarities) are represented by the four tests in the battery.

1) Used by Cronbach, L.J. and Snow, R.E. (1969, 64).

Imagine now that from an actual data set of TABLE 25 a correlation matrix is derived as exemplified in TABLE 26.

TABLE 26. Correlation matrix of a 2 x 2 factorial test design.

	11	12	21	22
11	1,00	0,78	0,42	0,32
12	0,78	1,00	0,28	0,38
21	0,42	0,28	1,00	0,82
22	0,32	0,38	0,82	1,00

As is well known from factor analytic reasoning, all information on the structuring of a set of standardized variables is contained in the correlation matrix. Such matrices may therefore be subjected to analytical treatments. While traditional methods of factor analysis mostly are performed on unstructured correlation matrices with a view to cluster variables a posteriori, the intention now is to analyze a correlation matrix that is structured a priori on a rational basis. Characteristic for the correlation matrix of TABLE 26 is that the tests making up the battery are already clustered into groups of variables. The purpose of an analytical approach to the structured correlation matrix should be to bring forth empirical evidence to show to what extent the rational factors can be corroborated.

Two approaches will be explicated in analyzing the correlation matrix above. These are extensions of the procedures for the eta and alpha analyses introduced in PART 1. The character of a rationalistic approach to factor analysis by way of ANOVA should become convincingly clear when dealing with systematic and balanced test designs as illustrated in TABLE 26.

3.1.1. The matching of linear combinations to test design.

Assuming the raw scores of TABLE 25 have been transformed to standard scores, the total variance of the test battery will be 4, which is equal to the trace of the correlation matrix of TABLE 26.

According to linear algebra, the total variance of the four tests can be given as the sum of four linear combinations of the four tests, provided the linear combinations are made orthogonal to each other.

In ANOVA, a factorial design on the variables in a repeated measures approach means that fixed linear combinations are employed. There are many ways of making four orthogonal linear combinations of four tests; yet ANOVA sticks to a routine way of combining. For the present test design the four routinely defined linear combinations are given by a design matrix that was anticipated in 2.2.4.d) and shown again in TABLE 27, where

TABLE 27. Design matrix for prestructured tests.

	11	12	21	22
I	+1	+1	+1	+1
II	+1	+1	−1	−1
III	+1	−1	+1	−1
IV	+1	−1	−1	+1

column headings denote the structured tests as explained for TABLE 25, roman numerals for rows denote the four linear combinations. The design matrix is also a weighting matrix, indicating that all tests get the same weight, although the combinations of signs are different for the four linear combinations. The design matrix of TABLE 27 is made up of four orthogonal linear combinations, shown by the fact that the productsum of each possible pairing of weights for two linear combinations at a time will always be zero.

In terms of the concrete tests in the example used, the first linear combination (I) is the sum score across all four tests. The three other linear combinations are difference scores, variously defined. The second one (II) is the contrast between the verbal and performance tests, the third (III) the contrast between opposites and similarities tests. The fourth linear combination is a much more complex difference score: It is the difference of two difference scores. The derivation of this particular linear combination makes its property more easily seen:

$$(\overset{11}{+1} - \overset{12}{1}) - (\overset{21}{+1} - \overset{22}{1}) = \overset{11}{+1} - \overset{12}{1} - \overset{21}{1} + \overset{22}{1} = (IV)$$

It is the variance of these four linear combinations that are of interest; each of them conveying information on how well the variously defined scores, based on rational considerations, can discriminate among individuals. The sum score variance indicates individual differences in a common trait across the system of four tests. The difference score between verbal and performance tests indicates to what extent some persons get high scores on verbal and low ones on performance tests, and vice versa for other persons. The difference score between opposites and similarities tests implies differential aptitudes in the two formats. While the three first linear combinations are fairly easy to interpret, the fourth one, the difference of differences score, is more difficult, though not impossible to interpret.

The variances of the four linear combinations, defined by the design matrix of TABLE 27, will exhaust the total test variance of the standardized battery and can be written the following way,

$$V_{I} = 1/(n-1)\Sigma(z_{11} + z_{12} + z_{21} + z_{22})^2 \quad (1)$$
$$V_{II} = 1/(n-1)\Sigma(z_{11} + z_{12} - z_{21} - z_{22})^2 \quad (2)$$
$$V_{III} = 1/(n-1)\Sigma(z_{11} - z_{12} + z_{21} - z_{22})^2 \quad (3)$$
$$V_{IV} = 1/(n-1)\Sigma(z_{11} - z_{12} - z_{21} + z_{22})^2 \quad (4)$$

where symbols are used as above. In addition, z symbolizes standard scores.

In expanding formulas (1), (2), (3), and (4), the variance structures of the four linear combinations will emerge. It can easily be seen that the expansions will result in veriance and covariance terms. As scores are standardized, the variances obtained must be unity, and the covariances correlation coefficients.

The correlation coefficients may conveniently and meaningfully be classified into categories, depending on the identifying aspects of the tests correlated. Three such categories of covariance can be established, each of them conveying information on common variance across combinations of tests:

a) The correlation coefficients between tests that are dissimilar both as regards trait (verbal/performance) and format (opposites/similarities). There are two different combinations of such heterotrait-heteroformat tests, 11/22 and 12/21. However, the obtained correlation coefficients are substantively equivalent: They bring forth what is common across maximally dissimilar tests, as far as the test design goes.

b) The correlation coefficients between tests that are dissimilar as regards trait, similar as regards format. There are two different combinations of such heterotrait/monoformat tests, 11/21 and 12/22. Again, it should be understood that the coefficients obtained bring forth the same kind of information as regards commonness among tests that are measuring different traits by the same formats.

c) The correlation coefficients between tests that are similar as regards trait, dissimilar as regards format. Two different combinations of such monotrait/heteroformat tests can be found, 11/12 and 21/22. Here, the correlation is about commonness among tests measuring the same trait by different formats.

The categories of covariance so defined should be identified in the correlation matrix of TABLE 26, and their substantive meaning made clear.

As noted, there are two different combinations of tests for each of the three categories of covariance identified. Let it be assumed that there is no reason to believe that the two combinations of tests within each category will result in significantly different correlation coefficients, so that each category of covariance might be represented by the average of the two correlation coefficients. According to the above classification of covariances the average coefficients will be denoted $\bar{r}_a$, $\bar{r}_b$, and $\bar{r}_c$.

By expanding formulas (1), (2), (3), and (4) and subsequent reduction, one obtains the following variances of the four linear combinations:

$$V_I = 4 + 4\bar{r}_c + 4\bar{r}_b + 4\bar{r}_a \quad (5)$$

$$V_{II} = 4 + 4\bar{r}_c - 4\bar{r}_b - 4\bar{r}_a \quad (6)$$

$$V_{III} = 4 - 4\bar{r}_c + 4\bar{r}_b - 4\bar{r}_a \quad (7)$$

$$V_{IV} = 4 - 4\bar{r}_c - 4\bar{r}_b + 4\bar{r}_a \quad (8)$$

It should be understood that formulas (5), (6), (7), and (8) are different ways of combining the elements in the correlation

TABLE 28. Correlation matrices for the linear combinations.

V_I

	11	12	21	22
11	+	+	+	+
12	+	+	+	+
21	+	+	+	+
22	+	+	+	+

V_{II}

	11	12	21	22
11	+	+	−	−
12	+	+	−	−
21	−	−	+	+
22	−	−	+	+

V_{III}

	11	12	21	22
11	+	−	+	−
12	−	+	−	+
21	+	−	+	−
22	−	+	−	+

V_{IV}

	11	12	21	22
11	+	−	−	+
12	−	+	+	−
21	−	+	+	−
22	+	−	−	+

matrix. Formula (5) is based on the far most common way of considering a correlation matrix, namely, the variance of the sum score across the four tests. But formulas (6), (7), and (8) are also based on quite legitimate versions of the correlation matrix, although not often paid attention to. The variances of the four linear combinations given as functions of the correlation matrix can be seen from TABLE 28, where only the signs of the elements are written, not the elements (correlation coefficients) themselves. The reader should be convinced that TABLE 28 follows from formulas (1), (2), (3), and (4).

While the upper-left matrix of TABLE 28 is the wellknown sum score variance, the upper-right matrix is the difference score variance when the sum score for performance tests is subtracted from the sum score of the verbal tests. As indicated above, this is a meaningful score. Similarly, the lower-left matrix is the difference score variance of opposites scores contrasted with similarities scores. The lower-right matrix gives the variance for the difference score of the opposites-similarities difference for the verbal test and the opposites-similarities difference for the performance test. The meaning of this complex score may be indicated by the fact that for one person the sign of the difference score between opposites and similarities for the verbal test may be reversed for the difference score between opposites and similarities for the performance test. As mentioned above, the full meaning of this triple interaction score is not easily grasped.

3.1.2. Eta analysis (Sums of squares analysis).

The four linear combinations thus constructed have a parallel in an ordinary ANOVA approach to the data matrix of TABLE 25, where the sources of variance associated with persons define the same linear combinations. TABLE 29 presents the ANOVA

TABLE 29. ANOVA of 2 x 2 factorial test design.

Lin. comb.	Source	SS	df	MS	Manifest variance structure
I	Persons	SS_p	(n-1)	MS_p	$= 1 + \bar{r}_c + \bar{r}_b + \bar{r}_a$
	Trait	SS_t	1	MS_t	$= 0$
	Format	SS_f	1	MS_f	$= 0$
	T x F	SS_{tf}	1	MS_{tf}	$= 0$
II	P x T	SS_{pt}	(n-1)	MS_{pt}	$= 1 + \bar{r}_c - \bar{r}_b - \bar{r}_a$
III	P x F	SS_{pf}	(n-1)	MS_{pf}	$= 1 - \bar{r}_c + \bar{r}_b - \bar{r}_a$
IV	PxTxF	SS_{ptf}	(n-1)	MS_{ptf}	$= 1 - \bar{r}_c - \bar{r}_b + \bar{r}_a$
	Total'	$SS_{tot'}$	4(n-1)	MS's = 4	

of the standardized scores for the n x 2 x 2 factorial test design. In standardizing the scores, zero variance is obtained for traits and formats. The same must consequently be the case for the trait by format interaction.

What needs explanation in TABLE 29 is the changed values of the variances of the four linear combinations (I), (II), (III), and (IV) expressed as MS_p, MS_{pt}, MS_{pf}, and MS_{ptf}, respectively, compared to the values of formulas (5), (6), (7), and (8). Obviously, the values in TABLE 29 are 1/4 of the obtained variances in formulas (5) - (8). This reflects the different ways of defining variances, as dealt with in 2.1.1.e). Those definitions may bear a repetition here.

The variances in formulas (5) - (8) are based on total scores of the linear combinations, either a sum or a difference. On the other hand, the variances in TABLE 29 (the manifest variance structures) are based on the average values of those linear combinations. When a sum is obtained by adding k test scores, the variance of this linear combination can be expressed in three ways; as a sum score variance (V_X), as a MS in ANOVA (MS_p), and as the variance of an average score ($V_{\bar{X}}$). The relationship between the three variances will be,

$$V_X = kMS_p = k^2 V_{\bar{X}}$$

By being reminded of these relations, it should be clear how the variance structures of the linear combinations in TABLE 29 have been reached.

What is evident from TABLE 29 is that the sum of the variances of the four linear combinations is 4 which is the total variance of the four standardized tests. In multivariate statistics this sum of variances is also called the trace, which here is the sum of the principal diagonal in the correlation matrix of TABLE 26. Thus, from TABLE 29 it can be seen that the total test variance is transformed to the sum of the variances of the four linear combinations. Here is where the parallel to factor analytic reasoning becomes apparent. The solution in TABLE 29 is a full rank solution in that the total test variance (the trace) has been accounted for completely by four othogonal linear combinations. Particular for the present solution is that unit weights are used in accordance with the design matrix presented in TABLE 27. Also unit variances are used in the principal diagonal of the correlation matrix, and not communalities, the point of departure for factor analysis. Thus, the present ANOVA approach is close to a centroid solution as traditionally conceived, but not identical, however. As noted in 2.2.2., there is a subtle difference connected with the ways ANOVA and FA account for variance. While FA accounts for variance by averaging the squared loadings, ANOVA does this by squaring the average loading.

In terms of the present procedure, the proportions of total variance accounted for by the linear combinations can be given as ratios of sums of squares.

$$Eta_I^2 = \frac{SS_p}{SS_{tot}} = \frac{MS_p(n-1)}{4(n-1)} = \frac{1 + \bar{r}_c + \bar{r}_b + \bar{r}_a}{4} \qquad (9)$$

$$Eta_{II}^2 = \frac{SS_{pt}}{SS_{tot}} = \frac{MS_{pt}(n-1)}{4(n-1)} = \frac{1 + \bar{r}_c - \bar{r}_b - \bar{r}_a}{4} \qquad (10)$$

$$ETA_{III}^2 = \frac{SS_{pf}}{SS_{tot}} = \frac{MS_{pf}(n-1)}{4(n-1)} = \frac{1 - \bar{r}_c + \bar{r}_b - \bar{r}_a}{4} \qquad (11)$$

$$Eta_{IV}^2 = \frac{SS_{ptf}}{SS_{tot}} = \frac{MS_{ptf}(n-1)}{4(n-1)} = \frac{1 - \bar{r}_c - \bar{r}_b + \bar{r}_a}{4} \qquad (12)$$

As sums of squares ratios, formulas (9), (10), (11), and (12) are squared correlation ratios, or eta squared coefficients. But the formulas are also traditional per cent trace indices, as they are ratios of the variance of the linear combinations to total test variance, which is trace.

In principle, the present approach in accounting for explained variance is not different from the per cent trace indices as established, say, by a principal components analysis. While a principal components analysis in turn maximises the variance of each linear combination by differential weighting, the ANOVA procedure as here adopted employs rational weights, equal in absolute values, but differing in signs according to a systematic structuring of tests, prior to the analysis. The linear combinations of substantive interest are just those generated by the design matrix.

From the variance accounting procedure described above in terms of squared eta coefficients, the structure of total test variance can be given as a composition of the squared etas. As the proportion contributed by each of the variances of the four linear combinations is based on a combination that is uncorrelated to the three other combinations, the eta squared coefficients can be added to yield a nonoverlapping accounting of the observed test variance. By setting total test variance to unity,

$$1 = \text{Eta}_{IV}^2 + \text{Eta}_{III}^2 + \text{Eta}_{II}^2 + \text{Eta}_{I}^2 \qquad (13)$$

The variance structure of (13) is a manifest structure in that the linear combinations are all on the observed level. No inference to underlying features in the tests is implied. The variances of each of the linear combinations given in terms of the properties of the correlation matrix, are also manifest structures obtained as compositions of the variances and covariances in the correlation matrix.

3.1.3. Alpha analysis (Variance components analysis).

In dealing with test scores it is not uncommon to go beyond the observed level to some underlying traits that are thought to be the elements constituting a latent structure of the observed scores. On an inferred level a theoretical structure can be imposed on the manifest variances of the different linear combinations. Such a structure will reflect the way one thinks the variances are generated in terms of particular contributions from each of the identified sources of variance in the test design. An intuitive logic of a latent variance structure of the variance-covariance matrix (or the correlation matrix) can be developed, indicating how the variance and the different categories of covariance previously defined in 3.1.1. are construed to be built up.

a) The logic of an intuitive approach to a latent structure.

Consider again the socalled heterotrait-heteroformat correlations in the matrix of TABLE 26. The two correlation coefficients constituting this particular category reflect the covarying of tests that are maximally dissimilar on a rational basis, as they are dissimilar in both of the two identifying aspects of the test design. Therefore, this category of covariance should be taken to be a measure of a common-to-all-four-tests variance. It should also be expected to yield the least correlation in the matrix. Let this base-line correlation be represented by the average heterotrait-heteroformat covariance. This average value will be called a covariance component, and symbolized ρ_a.

The heterotrait-monoformat type of correlation reflects a covariance that has something more in common than the heterotrait/heteroformat correlation. The tests correlated under this category are similar as to format, and are expected to have a common component because of this, as they are both either opposites or similarities tests. But the tests differ in being either a verbal or a performance test. It is reasonable to conceive of this heterotrait-monoformat correlation as consisting of two covariance components. First, the common-to-all-four-tests variance is naturally included as one of the components. What is common to tests that differ in two identifying aspects should also be common to tests that differ in only one of the identifying aspects. In addition, it is reasonable to conceive of a covariance component attributable to the commonness of the correlated tests being both either opposites or similarities tests. Let this component be called ρ_b. The conceptual structure of an average heterotrait-monoformat correlation can thus be given the form,

$$\bar{r}_b = \rho_b + \rho_a$$

The monotrait-heteroformat tests, in addition to the common-to-all-four-tests variance, in their correlation will also reflect a covariance component indicating what is specifically common when verbal test are correlated, and when performance tests are correlated, separately. Because of the orthogonal design, there will be no influence of the heterotrait-monoformat covariance proper on the monotrait-heteroformat correlation. The same is of course the case the other way around: No influence of the monotrait-heteroformat covariance proper on the heretotrait-monoformat correlation. According to this kind of reasoning, the conceptual structure of an average monotrait-heteroformat correlation will be, calling the component of concern now ρ_c,

$$\bar{r}_c = \rho_c + \rho_a$$

After having imposed the inferred covariance structure on the two categories of correlation, it remains to see what kind of structure the unit test variances might reasonably be given.

The belongingness among the four tests established by rationally knitting them together in a systematically constructed battery, makes it sensible to conceive of the single test variance as composed of the different covariance components so far defined, indicating the components of togetherness in this particular family of tests. It is also reasonable to conceive of a unique test variance component left over after the imposement of the three covariance components on the test variance. This residual will conceptually be a mixture of a test-specific component and an error component due to an assumed unreliability of the tests. This replication error can not be isolated in the present design since only one observation is obtained within each of the cells in the design.

By calling the unique component σ^2_{res}, the intuitive logic of how the latent structure of the test score variance may be composed, can now be written,

$$1 = \sigma^2_{res} + \rho_c + \rho_b + \rho_a \qquad (14)$$

It should be noted that (14) is the variance structure of of one (average) test; it will be the latent structure of the principal diagonal in the correlation matrix.

In accounting for the latent structure of the linear combinations, one can elaborate on (14) by bringing together the structural theory of the different categories of covariance and the variance in a latent structure of the complete correlation matrix. Then a fairly complex structure of the correlation matrix

TABLE 30. Latent structure of 2 x 2 correlation matrix.

	11	12	21	22
11	$\sigma^2_{res} + \rho_c + \rho_b + \rho_a$	$\rho_c + \rho_a$	$\rho_b + \rho_a$	ρ_a
12	$o^2_c + \rho_a$	$\sigma^2_{res} + \rho_c + \rho_b + \rho_a$	ρ_a	$\rho_b + \rho_a$
21	$\rho_b + \rho_a$	ρ_a	$\sigma^2_{res} + \rho_c + \rho_b + \rho_a$	$\rho_c + \rho_a$
22	ρ_a	$\rho_b + \rho_a$	$\rho_c + \rho_a$	$\sigma^2_{res} + \rho_c + \rho_b + \rho_a$

will emerge, as can be seen from TABLE 30, where the latent structure of the sum score variance is presented. The latent structure of the variance of the three other linear combinations would appear by applying the appropriate signs according to TABLE 28.

By reassembling similar components of TABLE 30 and summing, the correlation matrix can be written as a sum of weighted components. Thus, the sum score variance as a latent structure will be,

$$V_I = 4\sigma^2_{res} + 8\rho_c + 8\rho_b + 16\rho_a \qquad (15)$$

The sum score variance as given by (15), which is the sum of the correlation matrix, is 4 times larger than the MS_p, as noted previously. Therefore, the ANOVA result can be written,

$$MS_p = \sigma^2_{res} + 2\rho_c + 2\rho_b + 4\rho_a \qquad (16)$$

b) The formal ANOVA approach to a latent structure.

It is indeed interesting that the structure developed in (16) on a more intuitive basis can be obtained formally by employing rules of thumb in writing out the structural models (in traditional ANOVA, the E(MS's)) for the sources of concern in TABLE 29.

TABLE 31. Latent structure of linear combinations.

Lin. comb.	Source	Observed MS	Structural model for MS
I	MS_p	$= 1 + \bar{r}_c + \bar{r}_b + \bar{r}_a$	$= \sigma^2_{ptf} + 2\sigma^2_{pf} + 2\sigma^2_{pt} + 4\sigma^2_p$
II	MS_{pt}	$= 1 + \bar{r}_c - \bar{r}_b + \bar{r}_a$	$= \sigma^2_{ptf} + 2\sigma^2_{pt}$
III	MS_{pf}	$= 1 - \bar{r}_c + \bar{r}_b - \bar{r}_a$	$= \sigma^2_{ptf} + 2\sigma^2_{pf}$
IV	MS_{ptf}	$= 1 - \bar{r}_c - \bar{r}_b + \bar{r}_a$	$= \sigma^2_{ptf}$

The similarity in structure between (16) and the model for MS_p in TABLE 31 is evident. However, it remains to be seen how the components of (16) can be matched to the components of TABLE 29. This amounts to establishing the relations between the intuitively defined covariance components of (16) and the more traditional, and formally defined, variance components of TABLE 31.

It should be clear from the variance components model (the latent structure model) of TABLE 31, that the separate <u>variance</u> components can be derived in terms of average correlation coefficients. First, the weighted components will be:

$$\sigma^2_{ptf} = 1 - \bar{r}_c - \bar{r}_b + \bar{r}_a$$

$$2\sigma^2_{pf} = (1 - \bar{r}_c + \bar{r}_b - \bar{r}_a) - (1 - \bar{r}_c - \bar{r}_b + \bar{r}_a) = 2\bar{r}_b - 2\bar{r}_a$$

$$2\sigma^2_{pt} = (1 + \bar{r}_c - \bar{r}_b - \bar{r}_a) - (1 - \bar{r}_c - \bar{r}_b + \bar{r}_a) = 2\bar{r}_c - 2\bar{r}_a$$

$$4\sigma^2_{p} = (1 + \bar{r}_c + \bar{r}_b + \bar{r}_a) - (2\bar{r}_c - 2\bar{r}_a) - (2\bar{r}_b - 2\bar{r}_a) - (1 - \bar{r}_c - \bar{r}_b + \bar{r}_a) = 4\bar{r}_a$$

After this, the unweighted components can be defined by taking account of the proper coefficients in the structural models:

$$\sigma^2_{p} = \bar{r}_a \qquad (17)$$

$$\sigma^2_{pt} = \bar{r}_c - \bar{r}_a \qquad (18)$$

$$\sigma^2_{pf} = \bar{r}_b - \bar{r}_a \qquad (19)$$

$$\sigma^2_{ptf} = 1 - \bar{r}_a - (\bar{r}_c - \bar{r}_a) - (\bar{r}_b - \bar{r}_a) = 1 - \bar{r}_c - \bar{r}_b + \bar{r}_a \qquad (20)$$

(Here, σ^2_{ptf}, although already given from TABLE 29, has been derived as a residual variance component. By definition, it should be, $\sigma^2_{res} = \sigma^2_{ptf} = 1 - \sigma^2_{pf} - \sigma^2_{pt} - \sigma^2_{p}$.)

From (17), and in recalling the definition of ρ_a, the common covariance component, it is clear that $\rho_a = \sigma^2_p$. This is a corroboration of the result obtained by Eikeland (1970) concerning the general component in the ANOVA approach to the Hoyt-Cronbach coefficient alpha, indicating that it should rather be called a covariance component. As regards the variance components for the person by trait and the person by format interactions compared to the covariance components, ρ_c and ρ_b, a seemingly confusing result is obtained, in that it can be shown that,

$$\rho_b = \bar{r}_b - \bar{r}_a = \sigma^2_{pf}, \text{ and } \rho_c = \bar{r}_c - \bar{r}_a = \sigma^2_{pt}$$

What appears somewhat unreasonable at first sight is that one has to use the correlation between traits in order to define the person by format interaction component and the correlation between formats to define the person by trait interaction component. Intuitively, it is certainly not easy to see that after all this relationship is reasonable. The clue to an understanding is contained in the complementary character of the two central concepts in this context, namely covariance and interaction. They are complementary in the sense that the more we have of the one the less we have of the other, and vice versa.

The conception of a partitioning of the sum score variance into covariance components, like what is done in (16), is compatible with FA reasoning, while the partitioning by way of variance components is not. Nevertheless, the variance structures established either way are identical, and the ANOVA approach should be considered a genuine FA procedure, we think. However, one should remember that one linear combination's covariance component might turn up as another linear combination's variance component, when interactions (difference scores) are involved.

The results shown in (17), (18), (19), and (20) amounts to showing that the latent structures conceived of in terms of covariance coponents are completely convergent with the structural models established by ANOVA.

c) Alpha coefficients as indices of explained variances.

The sum of the four unweighted variance components in (17), (18), (19), and (20) is unity, which is the variance of one test,

$$1 = \sigma^2_{ptf} + \sigma^2_{pf} + \sigma^2_{pt} + \sigma^2_{p} \qquad (21)$$

This partitioning of the test variance of 1 into a sum of covariance components and a residual component, is a true counterpart to classical, Spearman-Thurstone FA tradition. In effect, what is being accomplished is to obtain an estimate of communality by way of the over all covariance component, σ^2_p, running through all of the tests, whatever the category; and the partially common components, σ^2_{pt} and σ^2_{pf}, running through some of the tests, but not all. In one respect the present system is more akin to Spearman than to Thurstone, in that the ANOVA structure is typically hierarchical, as can readily be observed from the latent structure representation of the correlation matrix in TABLE 30.

A variance accounting procedure in terms of components, either unweighted or weighted, is here considered an alpha analysis approach, as pointed out previously. Cronbach's alpha as originally conceived is seman-tically linked to reliability theory, although generalizability theory may have paved the way for a broader outlook. Syntactically, however, the alpha coefficient may certainly be looked upon as a very general construct, in essence reflecting the ratios of variance components.

According to this generic conception of alpha, all of the four components in (21) above can directly be interpreted as alpha coefficients, since the components are already given proportions. One could also be interested in knowing how much of the variance is accounted for by the covariance components together, like a communality measure,

$$\text{alpha}_{\text{com}(1)} = \frac{\sigma^2_{pf} + \sigma^2_{pt} + \sigma^2_{p}}{\sigma^2_{ptf} + \sigma^2_{pf} + \sigma^2_{pt} + \sigma^2_{p}} = \sigma^2_{pf} + \sigma^2_{pt} + \sigma^2_{p} \qquad (22)$$

The ratio of the sum of the common components to total variance in (22) constitutes an alpha coefficient according to the present conception.

While the latent structure of the variance of one average test as shown by (21) is conceived as a sum of unweighted components, the latent structure of the actual linear combinations of tests as revealed by the structural models for the MS's in TABLE 31 is considered a sum of differentially weighted components, the weights being a direct reflection of the test design.

For the test user it is of considerable interest to be able to interpret the scores of the linear combinations, substantively. The components of the latent structures are thought to bring forth information of some underlying traits that influence the test scores, whether a sum score or various difference scores. Each of the components conveys information of what is measured by total test scores, providing the test user with crucial cues for interpretation.

What is here at issue is a kind of parallel to the reliability of one test as contrasted to the reliability of the composite test. The structural model for the sum score variance in TABLE

31 (the MS_p) is to the author nothing less than an extension of the traditional Spearman-Brown prophecy formula for a homogeneous composite test to a test battery, say a differential aptitude test. A very general Spearman-Brown rationale seems to be formulable in terms of the latent variance structures as a composition of weighted variance components for whatever complex test design conceivable.

From the structural models of the linear combinations as given by TABLE 31, quite a few alpha coefficients can be constructed. Consider first the sum score variance, the MS_p. Each weighted component can here be given as a proportion by setting $MS_p = 1$. Each of them will be an alpha coefficient, the form of which can be exemplified by the coefficient for the first factor, the common factor, thought to be running through all of the tests:

$$\text{alpha}_{p(4)} = \frac{4\sigma^2_p}{\sigma^2_{ptf} + 2\sigma^2_{pf} + 2\sigma^2_{pt} + 4\sigma^2_p} \qquad (23)$$

While MS_p is the observed variance of the linear combination called the sum, formula (23) is an indication of how much of this observed variance is estimated to be explained by an underlying, common trait. By putting the weighted component for each of the three other sources in the structural model for MS_p as a numerator, three additional alpha coefficients are derived, which give the proportion of sum score variance explained by the other factors in the model. The additional alpha coefficients may be named $\text{alpha}_{pt(4)}$, $\text{alpha}_{pf(4)}$, and $\text{alpha}_{ptf(4)}$. (The subscript of 4 is indicating that a linear combination involving all four tests is being considered.) By summing all of the four alpha coefficients, total sum score variance, which is set to unity ($MS_p = 1$), is obtained,

$$MS_p = 1 = alpha_{ptf} + alpha_{pf} + alpha_{pt} + alpha_p \qquad (24)$$

In (24), $alpha_{ptf}$ will be the proportion of uniqueness in the sum score variance (specificity + error), and for most practical purposes interpreted as noise. The sum of the three other coefficient, however, will be a measure of signal in the system, in that the sum signifies the proportion of sum score variance that can be interpreted: $Alpha_{pf}$ indicates to what extent opposites tests are measuring other things than similarities tests, and $alpha_{pt}$ the extent to which verbal and performance tests are measuring distinct traits. $Alpha_p$, as mentioned, is a measure of how much common-to-all-four-tests variance is represented in the observed sums score variance. Recalling the relationships established in 3.1.3.b) between the covariance components for the two simple interactions and the variance components for the same interaction terms, $\rho_c = \sigma^2_{pt}$, and $\rho_b = \sigma^2_{pf}$, one can interpret $alpha_{pf}$ as a measure of how much of the sum score variance is explained by the commonness between traits (verbal and performance), and $alpha_{pt}$ as a measure of how much is accounted for by the commonness between the two formats, opposites and similarities.

Thus, according to (24) we are able to judge to what extent the sum score is interpretable. A meaningful use of that score seems to be dependent on a great contribution by $alpha_p$, and minimal contributions by the other coefficients.

In case of considerable interaction effects, the factor structure of the test battery will be such that it might be more meaningful to pay attention to difference scores rather than to the sum score. Two such difference scores are here at issue: The difference score between traits (the difference score between the verbal and performance tests), and the difference score between formats (the difference score between opposites and similarities tests). The observed variances of these difference scores are the MS_{pt} and the MS_{pf}, respectively. As can be seen from TABLE 31, the latent trait model for these scores are represented by the structural equations of variance components. This means that one can find how much of observed difference score variance is explained by a genuine (true) interaction component. In more traditional FA language this interaction component would probably match a bipolar factor.

In a strict test theoretical context, the structural models for MS_{pt} and MS_{pf} would be the point of departure for estimating the reliability of the difference scores. (A paper on the difference score issue is being prepared by the author, where the problems touched here will be dealt with more thoroughly.)

3.2. A worked-out numerical example.

We are now ready for an analysis of the correlation matrix of TABLE 26. In keeping with the rationales developed, two principally different approaches will be examined.

First, we are going to analyze the manifest or observed correlation matrix by way of a sums of squares analysis, which is called an eta analysis. The second approach will be based on the latent correlation matrix, or the inferred structure imposed on the correlation matrix as a theoretical construction of how the matrix may be thought to be generated in terms of the sources of variance going into the test design. This latent structure analysis we called an alpha analysis. Two problems will be faced in the alpha analysis, (1) how to account for the variance structure of one average test, and (2) how to account for the variance structure of the linear combinations of the test scores.

3.2.1. Eta analysis of the correlation matrix.

The partitioning of total test variance of the four standardized tests (which is 4) into variances of the four orthogonal linear combinations as reflected in a 2 x 2 factorial test design, can be performed by just summing coefficients in the intertest correlation matrix by taking into account the appropriate signs determined by the particular linear combinations.

The sign matrices to be used are found in TABLE 28. By combining signs with correlation coefficients and summing, the following variances of the linear combinations are directly obtainable:

$V_I = V_p = 10,0$

$V_{II} = V_{pt} = 4,4$

$V_{III} = V_{pf} = 1,2$

$V_{IV} = V_{ptf} = 0,4$

Total $= 16,0$

By taking ratios of the variances for each of the linear combinations to total variance, squared eta coefficients would be obtained, indicating the proportion of total variance accounted for by the manifest factors.

Instead of doing this, one can do an ordinary ANOVA of the standardized data matrix upon which the correlation matrix of TABLE 26 is based. If that data matrix were available, the following ANOVA table would be obtained. (As no original data matrix for the correlation matrix of TABLE 26 exists, values for the sums of squares and df involving persons can not be given. But, the MS's can be given on the basis of the correlation matrix only.)

TABLE 32. ANOVA of standard data matrix.

Source	Lin comb	SS	df	MS	Eta^2
Person	II	$MS_p(n-1)$	$(n-1)$	$MS_p = 2,5$	0,625
Trait		0	1	0	
Format		0	1	0	
T x F		0	1	0	
P x T	II	$MS_{pt}(n-1)$	$(n-1)$	$MS_{pt} = 1,1$	0,275
P x F	III	$MS_{pf}(n-1)$	$(n-1)$	$MS_{pf} = 0,3$	0,075
PxTxF	IV	$MS_{ptf}(n-1)$	$(n-1)$	$MS_{ptf} = 0,1$	0,025
Total'		$SS_{t'}$	$df_{t'} = 4(n-1)$	Trace $= 4,0$	1,000

The MS's in TABLE 32 are found by way of the formulas for manifest variance structures in TABLE 29. They are 1/4 of the variances obtained by summing the correlation matrix for the different linear combination. This is in keeping with the relationships established between the various ways of defining the variance of linear combinations, as shown in 2.1.1.e).

In TABLE 32 the sum of the MS's for the four linear combinations, each multiplied by degrees of freedom, (n-1), will be total sum of squares, here symbolized as SS_{t}. It should therefore be clear that the eta squared coefficients obtained can be computed by taking sums of squares ratios, compatible with a definition of an eta squared coefficient; although the most direct way of getting out the squared etas from TABLE 32 is to find the proportion of the MS,s of the linear combinations to trace.

The result of the eta analysis shows that 62 % of total test variance is sum score variance, 28% variance attributable to the difference score between verbal and performance tests, 8% to the difference score between opposites and similarities tests, and lastly, 2% of the total test variance is explained by the triple interaction, P X T X F, which in effect is the variance of a difference of differences score. How that linear combination can be interpreted this way, was shown in 3.1.1. on page 51.

3.2.2. Alpha analysis of the correlation matrix.

First, the unweighted variance components analysis is shown. According to TABLE 31, the structural models for the MS's, the components can be found by solving for the unknowns from the bottom of the table upwards. Both unweighted components and the MS's written as sums of weighted components are given in TABLE

TABLE 33. Weighted and unweighted variance components.

Source	MS	Composition of MS's	Unweighted components
MS_p	= 2,5	= 0,1 + 2. 0,1 + 2. 0,5 + 4. 0,3	σ^2_p = 0,3
MS_{pt}	= 1,1	= 0,1 + 2. 0,5	σ^2_{pt} = 0,5
MS_{pf}	= 0,3	= 0,1 + 2. 0,1	σ^2_{pf} = 0,1
MS_{ptf}	= 0,1	= 0,1	σ^2_{ptf} = 0,1
Trace = 4,0			V_t = 1,0

33. The sum of the unweighted components should add to unit test variance (V_t = 1). In the present analysis the unweighted components are directly interpretable as per cent explained variance, as the absolute components values are also proportions. The structure revealed is the latent structure of one average test variance. A common factor running through the whole battery, σ^2_p, accounts for 30% of the variance of one test. The person by trait interaction, σ^2_{pt}, explains as much as 50% of the variance. This means that a relatively large proportion of the variance can be attributed to the fact that persons are accomplishing differently on verbal and performance tests. This should be taken as an indication that verbal and performance tests are tapping different traits. On the other hand, the interaction between person and format indicates that opposites and similarities tests to a great extent are measuring the same thing: The variance component, σ^2_{pf}, is small, therefore a relatively great correlation between

formats exists. As a matter of fact, the average correlation between traits, and the average correlation between formats, can be reconstructed from the components. They will be, respectively,

$$\bar{r}_b = \sigma^2_{pf} + \sigma^2_p = 0,1 + 0,3 = 0,4$$

$$\bar{r}_c = \sigma^2_{pt} + \sigma^2_p = 0,5 + 0,3 = 0,8$$

This result can be confirmed by examining the correlation matrix, TABLE 26. It follows from the discussion in 3.1.3.a) and b).

Next, a weighted variance components analysis will be shown for the linear combinations, the sum score and the two difference scores.

In considering the sum score variance in terms of an inferred structure, the model for MS_p in TABLE 31 should be employed. The result can be taken from TABLE 33:

$$\begin{aligned} MS_p &= \sigma^2_{ptf} + 2\sigma^2_{pf} + 2\sigma^2_{pt} + 4\sigma^2_p \\ &= 0,1 + 2.\ 0,1 + 2.\ 0,5 + 4.\ 0,3 \\ &= 0,1 + 0,2 + 1,0 + 1,2 \\ 1,00 &= 0,04 + 0,08 + 0,40 + 0,48 \end{aligned}$$

The composition of the sum score variance given as a sum of proportions (setting MS_p = 1,00) shows that 48% of that variance is estimated to be contributed by a common factor, 40% attributable to a bipolar factor, indicating that verbal and performance tests are measuring different traits, and 8% of the variance is accounted for by another bipolar factor, the contrast between opposites and similarities tests.

Substantively considered, the result indicates that the sum score is hardly interpretable. This follows from the great contribution of the person by trait interaction. The advice recommended by the latent structure analysis of the sum score variance would probably be that verbal and performance tests should be kept separate, because they seem to represent distinct traits.

The latent structure of the person by trait interaction, the MS_{pt}, corroborates the above conclusion concerning the sum score:

$$MS_{pt} = \sigma^2_{ptf} + 2\sigma^2_{pt}$$
$$= 0,1 + 2 \cdot 0,5 = 0,1 + 1,0$$
$$1,00 = 0,09 + 0,91$$

As mentioned previously, in a strict test theoretical context the structure revealed by the model for MS_{pt} amounts to an estimation of the reliability of the difference score between verbal and performance tests. That estimate is extremely high and indicates even in an FA context that the two traits are fairly independent. (The reader should be reminded that the results discussed are hypothetical.)

A similar analysis could be made for the person by format interaction, the MS_{pf}. How that result would come out may be hinted at by referring to TABLE 33.

4. DISCUSSION

Some of the issues dealt with in this monograph have been discussed in the literature by Burt (1940), Burt (1947), Creasy (1957), and Bock (1960). Burt's (1940) treatment is especially relevant to Part 2, while Burt (1947) and the papers by Creasy and Bock are relevant to Part 3. Gollob (1968) is a very general discussion of the relationship between ANOVA and FA, but he does not bring in facets as part of the test design.

As can be seen from the references, no intensive discussion of the convergence of ANOVA on FA has taken place, and the papers are scattered over a long period. Characteristic for the lack of interest in the problem is that Burt's paper published in 1947 was reprinted almost 20 years later in Handbook of Multivariate Experimental Psychology, edited by Cattell (1966).

This state of affairs ought to be changed, because research workers are in need of a factor analysis approach that is more based on theory and a priori considerations as regards substance. To a very great extent traditional FA has been a posteriori in many respects, bordering on a very naive empiricism. The advantage of ANOVA as an a priori FA becomes especially evident when facets are introduced as part of the test designs, so that multiple dimensions are involved already in the theoretical stage of the research process, to be tested in the analysis of empirical data.

In the present effort to move somewhat forward in the direction of establishing a rationalistic alternative to purely a posteriori approaches to FA, the intention has been to clear the way toward an understanding by a relatively concrete discussion of the convergence of central concepts in the tradition of ANOVA on central concepts in the FA tradition.

One important objective has been to relate the manifest and latent covariance structures to the two models in the FA tradition: data reduction vs. FA proper, or actual (observed) vs. hypothetical (inferred) factors. In the literature, Burt (1940), Burt (1947), and Creasy (1957) deal with what is here called manifest covariance structures, while Peck (1960) is concerned with latent covariance structures. The present treatment has explicitly contrasted the two approaches, and made a clear distinction between the eta approach (manifest structures analysis) and the alpha approach (latent structures analysis). In addition, it is pointed out that the latent structures analysis can be both an average test variance decomposition and a sum score (or various difference scores) variance decomposition. The decomposition of the composite scores of the linear combinations into a weighted sum of variance components is believed to be an extension of traditional FA thinking, in that FA used to be concerned with the structuring of one (average) test's variance. For practical purposes it is obviously of great importance to be able to examine the variance structure of the linear combinations in actual use, in order to see whether the scores are meaningful, or to what extent they are interpretable.

The distinctions made are important to observe, we think. The manifest structure analysis, or the eta approach, deals only with the observed correlation matrices for the four orthogonal linear combinations of the four tests in the hypothetical data set used. The separate sums for the four matrices are the variances of the linear combinations, and the sum of these matrices exhausts the total test variance. The ratios of the separate matrix sums to total test variance are squared correlation ratios or eta squared coefficients, since these ratios can be shown also to be ratios on the sums of squares level. By definition, eta squared coefficients are sums of squares ratios.

No theory is implied, nor is any inference made in the manifest structure analysis. On the other hand, the latent structure analysis reveals an inferred variance structure of how the several sources of variance going into the test design are construed to influence the observed variances of the linear combinations. One is here interested in the influence of underlying traits or factors. While it should be acknowledged that a latent structure means that one is making an "inferential leap", that structure certainly provides a conceptual framework for making more challenging interpretations of what test scores may mean.

It has been argued in this monograph that ANOVA and FA are basically convergent relational systems. Nevertheless, there are conventional ways of doing things in the two traditions that are divergent. One divergence is particularly important to note. It concerns the way the proportions of total variance of the linear combinations are accounted for. In traditional FA the factor loadings are squared, summed, and then averaged. In the ANOVA approach the average loading is squared to get an eta squared coefficient. The interesting thing here is that the eta squared coefficient multiplied with total test variance yields the correct variance of the actual linear combination, while the FA procedure does not. The reader is referred back to 2.2.4.a) and b) as an example of this.

The three ANOVA ways of approaching a test score's interpretation by way of the structures of a correlation matrix, manifest and latent, may be in need of some concluding comments. In referring to the complex test design used as an example, we remember that all of the three approaches result in a structure consisting of four terms. It should be emphasized that these terms are not quite the same for the manifest and the latent structures.

In review, the three results, given in proportions, will be considered simultaneously:

(1)	Manifest total test variance	$0{,}025 + 0{,}075 + 0{,}275 + 0{,}625 = 1{,}000$ IV III II I
(2)	Latent average test score variance	$0{,}100 + 0{,}100 + 0{,}500 + 0{,}300 = 1{,}000$ σ^2_{ptf} σ^2_{pf} σ^2_{pt} σ^2_{p}
(3)	Latent sum score variance	$0{,}040 + 0{,}080 + 0{,}400 + 0{,}480 = 1{,}000$ σ^2_{ptf} $2\sigma^2_{pf}$ $2\sigma^2_{pt}$ $4\sigma^2_{p}$

(1) The proportions as brought forth by the manifest structure analysis, are the proportions of total test variance, i.e., the variance of the whole standardized data matrix, explained by the four observed orthogonal linear combinations of the four tests, one the sum score (I), the three others various difference scores (II, III, and IV). The sum of the variances of these linear combinations exhaust total test variance, which for the four standardized tests equals 4. The structure of (1) says

nothing about the internal structure of the scores of the linear combinations. The manifest structures of the variances of the observed scores obtained by forming the linear combinations can be seen from formulas (5), (6), (7), and (8) on page 54, or from TABLE 29. These are structures that are only functions of observed correlations among tests, and the test variances. The structures do not seem too interesting as a basis for interpreting the scores of the linear combinations.

The structure of the manifest total test variance as revealed by an eta analysis, provides no insight within the scores; it shows only the relative contribution by each of the linear combinations. A manifest structure analysis is explicit about this: No guesses are made of what might be the internal structure of test scores. However, that is the concern of a latent structure analysis.

(2) The proportions of explained variance of one average test score are the proportions conceived to be contributed by each of the unweighted components. The components reflect the elements thought to go into the structure of one average test score, indicating the contributions made to one single test's variance by the sources of variance identified in the test design. In the principal diagonal of TABLE 30, the conceptual structure of that variance is made concrete. It should be clear that such structures can not be meaningful without bearing in mind that the tests have been rationally tied together in a system defined by the test design. Thus, the tests are made to belong together. One test considered in isolation has no variance structure, but it certainly has a variance. The struc-

true property of a test variance can only be meaningful in the context of some defined belongingness among tests.

(3) The structural property of combinations of tests in terms of components seems intuitively more meaningful than the structure of one test. Also, the combination of tests is what test users employ in measuring individual differences. By far the most common linear combination is the sum score. The variance structure of such a score may be of crucial importance in interpreting what is measured by the sum score in terms of a common trait across all tests in a battery, and less common traits across subclusters of tests.

The proportions of explained sum score variance are the proportions contributed by each of the four weighted components. Here, the intention is to venture an inference of how much the different sources of test variance explain. This inference concerns a common trait running through all of the four tests regardless of classification, two less general components that explain common variance across traits (verbal-performance) and across formats (opposites-similarities), and lastly, a unique component.

In the latent structure analysis of (2) above, the average test score has got a theoretical structure, defined as nonoverlapping contributions by each of the four sources of variance in the test design (the sources involving individual differences). In the latent structure of (3), the sum score variance is conceived to be a function of the four variance components by a differential weighting system, depending on the number of levels of the different facets going into the test design.

that weighting system indicates, we think, the form the general Spearman-Brown rationale may take, when complex test designs are used. Accounting for the internal structure of a sum score variance has been the concern of test theory only. FA should also have good reasons for being interested in that structure. As far as the author knows, no FA technique has ever reached so far. The way ANOVA is used in this monograph strongly indicates that it should be possible to take advantage of the Spearman-Brown rationale also in an FA context.

The correlation matrix for the complex test design of TABLE 26 has deliberately been constructed so as to exaggerate a feature of the matrix that tends to make the manifest and the latent structure analyses quite different as to results. This feature is reflected in the fairly high correlation between opposites and similarities tests, and the relatively low correlation between verbal and performance tests. In an ANOVA context, the high correlation of opposites and similarities tests signifies a small person by format interaction, while the low correlation of verbal and performance tests signifies a large person by trait interaction.

In the observed sum score variance also the interactions are contributing sources without being identified. In the latent structure analysis of the sum score variance the different sources are distinguished so as to explain what kind of variance is going into the observed sum score variance.

Because of the great person by trait interaction, the latent structure analysis, both for one average test score and for the sum score, comes out with a result that the battery is not homogeneous enough to make a sum score substantively meaningful. This will not be discovered by an eta analysis, but an alpha analysis will. It should be emphasized that although the latent structure is an unobserved, theoretical structure, it is not without basis in some hard facts of what is going on in empirical data. It can be shown by manipulating the person by trait interaction (to give more or less interaction of this kind), that it directly influences the sum score variance. It should also be evident on a common-sense basis that this interaction will so influence sum scores, that the trait profiles for equal sum scores will be different. Then, certainly the meaning of the sum score gets blurred.

It is probably a good advice not to try to compare the results obtained by the three ways of analysis as discussed here, because they are hardly comparable. This will become the more convincing the more the basic rationales for the three approaches are being comprehended.

For pedagogical reasons our discussion has been deliberately concrete, tied to a particular test design. It should be understood that the procedure here described generalizes to any test design that is systematic and balanced. Such a priori test designs are believed to be very powerful instruments in making efforts to assess convergent and discriminant validities of constructs being introduced in differential psychology. They ought to be encouraged as a much needed counterbalance to a naive empiricism of an a posteriori character so often met in traditional FA.

5. REFERENCES

Bock, R. D. 1960. Components of variance analysis as a structural and discriminal analysis for psychological tests. The British Journal of Statistical Psychology, 8, 151-163.

Burt, C. 1940. The Factors of the Mind. London:University of London Press.

Burt, C. 1947. A comparison of factor analysis and analysis of variance. British Journal of Psychology, Statistical Section 1, 3-26.

Burt, C. 1966. A comparison of factor analysis and analysis of variance. In Cattell, R. B. (Editor). Handbook of Multivariate Psychology. Chicago: Rand McNally.

Cohen, J. 1968. Multiple regression as a general data-analytic system. Psychological Bulletin, 70, 426-443.

Creasy, M. A. 1957. Analysis of variance as an alternative to factor analysis. Journal of Royal Statistical Society, Series B, 19, 318-325.

** Eikeland, H.M. 1970. Coefficient alpha and the expected variance-covariance matrix of random composite measurements. Oslo. Mimeographed.

Eikeland, H. M. 1971a. On the generality of univariate eta. Scandinavian Journal of Educational Research, 15, 149-167.

Eikeland, H. M. 1971b. Correlational analyses of school marks influenced by multiple sources of variance. Oslo. Mimeographed.

Fennessey, J. 1968. The general linear model: A new perspective on some familiar topics. American Journal of Sociology, 74, 1-27.

Gollob, H. F. 1968. A statistical model which combines features of factor analytic and analysis of variance techniques. Psychometrika, 33, 73-116.

Gulliksen, H. 1950. *Theory of Mental Tests*. New York: John Wiley & Sons.

Harman, H. H. 1967. *Modern Factor Analysis*. (2nd edition). Chicago: University of Chicago Press.

Jennings, E. 1967. Mixed effects analysis of variance by regression analysis. *Multivariate Behavioral Research*, 2, 95-103.

* Lord, F. M. and Novick, M. R. 1968. *Statistical Theories of Mental Test Scores*. Reading, Mass.: Addison-Wesley.

Morrison, D. F. 1967. *Multivariate Statistical Methods*. New York: McGraw-Hill.

Nunnally, J. C. 1967. *Psychometric Theory*. New York: McGraw-Hill.

Rozeboom, W. W. 1966. *Foundations of the Theory of Prediction*. Homewood, Ill.: Dorsey Press.

Rulon, P. J. 1939. A simplified procedure for determining the reliability of a test by split-halves. *Harvard Educational Review*, 9, 99-109.

Tatsuoka, M. M. and Tiedeman, D. V. 1963. Statistical techniques. In Gage, N. L. (Editor). *Handbook of Research on Teaching*. Chicago: Rand McNally.

Winer, B. J. 1962. *Statistical Principles in Experimental Design*. New York: McGraw-Hill.

*Kennedy, J. J. 1970. The eta coefficient in complex ANOVA designs. *Educational and Psychological Measurement*, 30, 885-889.

**Cronbach, L. J. and Snow, R. E. 1969. Individual Differences in Learning Ability as a Function of Instructional Variables. Final Report. ERIC Document Reproduction Service ED-029-001.