6

COEFFICIENT ALPHA

AND

THE EXPECTED VARIANCE-COVARIANCE MATRIX

OF

RANDOM COMPOSITE MEASUREMENTS

Hans-Magne Eikeland

University of Oslo

Oslo, October 1970

Coefficient Alpha and the Expected Variance-Covariance Matrix of Random Composite Measurements

Hans-Magne Eikeland, University of Oslo[1)]

When coefficient alpha was introduced by Cronbach (1951), it was as a special case regarded as an exact determination of a particularly defined reliability coefficient, the coefficient of equivalence. However, coefficient alpha more generally conceived was described as an internal structure measure of tests, emphasizing a more independent role of alpha. Also, it was pointed out that alpha could be considered an estimate of the expected correlation between two tests drawn at random from a pool of items, a universe of items, thus anticipating more recent developments within test theory.

Nevertheless, coefficient alpha has continued to be conceived of mainly as a computing form of the equivalence coefficient and has mostly been derived as such, or as a lower bound estimate of reliability.

Though describing and interpreting alpha, Cronbach (1951) did not derive the coefficient. The derivation of alpha as an internal consistency measure per se appears to have attracted little attention in the test literature. With a conceptually clear separation between coefficient alpha and the generalizability coefficient as conceived by Cronbach, Rajaratnam, and Gleser (1963) or a more generally defined coefficient as conceived by Novick and Lewis (1967), it seems appropriate that one should be concerned with deriving alpha as an independent construct, standing on its own.

Within classical test theory coefficient alpha in its special KR 20 form or in its general form as adopted by Cronbach (1951) has been derived as a computing form for the parallel-form reliability by correlating rationally equivalent tests. Notable derivations within this tradition are Kuder and Richardson (1937), Jackson and Ferguson (1941), Gulliksen (1950), and Tryon (1957). In this interclass correlation type of derivation the concept of covariance has been a central one in trying to develop and explicate a rationale for coefficient alpha.

Alpha has also been derived by an analysis of variance technique, which represents an intraclass correlation type of derivation. In the intraclass correlation approach we apparently do not need the concept of rationally equivalent tests in deriving the coefficient. The first derivation of this kind within classical test theory model was made by Hoyt (1941), but he did not elaborate on alpha as an internal consistency measure.

Within the framework of a domain-sampling test theory model an analysis of variance rationale was adopted by Cronbach, Rajaratnam, and Gleser (1963) in deriving alpha as a lower bound estimate of the defined generalizability coefficient. Their derivation, although not easily grasped by the unsophisticated reader, represents a deep penetration into the logic and mathematics of alpha.

In the analysis of variance approach the concept of covariance, so useful in getting a feel for understanding the meaning of alpha, is no longer apparent. Alpha derived as an intraclass correlation coefficient has become somewhat obscure, probably by applying a technique of which the underlying rationale may not be readily understood, and also by using concepts that are not easily seen to be related to, or how related to, covariance. For example, in the interclass correlation approach we usually explicitly show that observed-score covariance equals universe-score variance, while in the intraclass correlation approach we are apt to lose sight of this aspect by directly making use of concepts like true-score variance, universe-score variance, or general component variance without showing the linkage to the covariance concept in this particular technical context.

The purpose of the present paper is to describe an intuitive and formal approach to alpha that makes explicit use of a covariance rationale in deriving the coefficient as an internal consistency construct within the framework of a domain-sampling test theory model. Particularly, the intention is to rederive alpha as a generalization of Tryon's (1957) recommended form of the reliability of an unstratified composite without resort to correlating rationally equivalent composites and where the full consequence of the random-sampling model is allowed for. Furthermore, the purpose is to show that the covariance approach as here adopted is formally equivalent to the analysis of variance approach. Particularly, it will be shown that what is called the general component variance in the analysis of variance approach is just another name for the expected covariance among items.

Essentially, nothing new is presented as the paper may be said to draw considerably upon Tryon (1957) and Cronbach, Rajaratnam, and Gleser (1963). Ho^wever, it is believed that the discussion will help to make clear how an alpha construct may be defined within a full-fledged random-sampling model and how conceived of in terms of an inferred structural property of the variance-covariance matrix. This basic structure is more or less hidden in the analysis of variance form, and also in the traditionally used computing form of alpha.

Tryon's derivation of alpha.

In classical test theory one frequently defines reliability in terms of covariance. As is well known, a product-moment correlation can be written,

$$\rho(x,y) = \frac{\sigma(x,y)}{\sigma(x)\sigma(y)} , \qquad (1)$$

which describes the correlation between x and y as the ratio of the covariance between x and y to the product of the two standard deviations. When x and y are conceived of as parallel tests, according to definition $\sigma(x)$ and $\sigma(y)$ are equal and the correlation between the two tests becomes equal to the reliability of one of the tests,

$$\rho(x,y) = \rho(x,x') = \frac{\sigma(x,x')}{\sigma(x)^2} , \qquad (2)$$

where the numerator is the sum of covariances among test samples[2)] from parallel tests.

Tryon (1957) derived his general form of the reliability of an unstratified composite, which equals alpha[3)], as a variant of (2) where under random sampling assumptions the numerator may be written as a function of the average covariance among test samples in the given test,

$$\rho(x,x') = \text{alpha}(k) = \frac{k^2\bar{c}(i,j)}{V(x)}, \text{where } i \neq j \quad (3)$$

In the domain-sampling model Tryon correlated two randomly drawn sets of test samples,each set consisting of k test samples,with no one-to-one correspondence between test samples in the two tests.Tryon thus changed from an item-parallel model,or what might be called a fixed-parallel model,to a random-parallel model which at first seemed to solve a long-standing estimation problem in the numerator of (2).

In the fixed-parallel model there are conceptually two types of covariance among test samples,one among fixed-parallel test samples,the other among random-parallel test samples.The first type of covariance is naturally enough assumed to be larger than the second type.As the interclass correlation approach to the equivalence coefficient presumes one given test and another hypothetical test parallel to the given one,there is no way of determining the covariance among fixed-parallel test samples as there exists no such pair of test samples.

While the estimation problem in the numerator of (2) within a fixed-parallel model only could be solved by making rather restrictive and unreasonable assumptions,as may be said did Kuder and Richardson (1937),Jackson and Ferguson (1941),and Gulliksen (1950),Tryon very elegantly came around the problem by his random-sampling model where there is only one type of covariance,the covariances among random-parallel test samples,which are determinable from the test samples within the given test.

On the other hand,Tryon in adopting his model ran into a difficult problem concerning the denominator in (2),which had caused no trouble in the classical model.In a full-fledged random-sampling model test variances are not necessarily equal.Consequently,nor does the product of the two standard deviations necessarily equal thetest variance.In order to escape this dilemma Tryon had to restrict his universe of random-parallel composites to sets of k randomly drawn test samples with equal variances and intercorrelations. Thus,Tryon ends up with a kind of Procrustean solution,and he did not succeed in having a break-through in test theory.

Another solution to Tryon's problem.

Now, Tryon's solution might have taken a somewhat different direction by allowing for unequal composite variances as the model actually requires. Adopting a random-sampling model and meeting the requirements of it, one may intuit an expected value of the variances of all random-parallel composites in the defined universe, all with k test samples.

Correlating two rationally random-parallel composites would give one an expected correlation coefficient where the numerator is the sum of all k^2 covariances, which all are the same expected test sample covariance, and where the denominator is the expected composite variance, being the product of two expected composite standard deviations.

Thus, instead of determining alpha as an exact correlation among comparable constructs in a Tryon sense, one can conceive of alpha as an expected correlation among random-parallel composites of k test samples each,

$$E(\rho(x,x')) = E(alpha(k)) = \frac{k^2 E(\sigma(i,j))}{E(\sigma(x)^2)} \qquad (4)$$

In (4) alpha is derived as an expected interclass correlation and is in accordance with one of Cronbach's (1951) interpretations of the coefficient.

However, what is now at issue, wanting to regard (4) as a pure internal consistency estimate, is to construct (4) as such by conceiving of a selfsufficient internal structure rationale. We want to derive alpha as an expected intraclass correlation where the concept of covariance is explicitly accounted for.

Imposing structure on the expected test sample variance.

In using tests, or measuring devices generally, one is mostly interested in observing individual differences. Reliability studies are mainly concerned with establishing statistical evidence that observed test variance is descriptive of real differences in traits measured, and not attributable to random and irrelevant sources of variation.

Having obtained only a single score by taking just one test sample observation on a sample of persons, there is no empirical basis to tell to what extent the distribution of this score represents genuine trait differences among the persons tested. Certainly, one can believe on a pure subjective basis, as for example the test maker naturally will do, that this single score distribution is descriptive of individual differences in the trait intended to measure. But this belief can not be verified, nor falsified, on statistical evidence. The belief reflects what Kelley (1942) called an a priori act of judgment on the part of the test maker, a prior conception of a construct to be checked by a quantitative statement when enough data are gathered.

In order to make the quantitative statement on which to base a conviction that one single score measure something or nothing, one has to administer at least two test samples believed to cover the same construct. To the extent that these repeated measurements covary, one is according to conventional practice provided with some statistical evidence that the multiple measurement taps some common element across the test samples administered, thus reinforcing the a priori act of judgment that went into the construction of the test samples.

The statistical evidence thus obtained is the only basis on which one can conclude that one test sample alone measures something systematically. The statement is conditional in that one has to specify to which defined class of observations this particular test sample belongs. It is only in relation to other test sample scores that one can have some assurance that a single test sample accounts for systematic variation. Conceiving of <u>belongingness</u> is the only way to give a test sample a frame of reference, and it leads naturally to the concept of a universe of observations. This is a crucial concept in generalizability theory as presented by Cronbach, Rajaratnam, and Gleser (1963). A defined universe of observations specifies what can be considered similar measures a priori. A test of the coherence in this universe is the covariance among randomly drawn test samples from this universe.

From the covariance among test samples one believes to have a basis for concluding to what extent one test sample measures a trait that the other test samples also measure. One specific test sample observation is probably of no particular interest. Rather, one tends to be interested in seeing to what extent one can expect whatever test sample is drawn to share a common variance with all other similarly defined test samples.

The average test sample covariance in the universe of observations, a pure conceptual value, is also the value expected between two actually drawn test samples. This expected covariance is now imposed on the expected test sample variance to indicate the capability of one average test sample observation to tap a trait common to all test samples within a defined universe of observations. The expected test sample variance is conceptually the average test sample variance in the universe. Both the expected covariance among test samples and the expected test sample variance are estimable from randomly drawn test samples.

By imposing the expected common variance shared by all test samples in the universe on the expected test sample variance, it is clear that one ignores test sample covariances shared by some test samples but not all. Also, it is apparent that one ignores the variance specific to each test sample. The common-to-some-test-samples variance and the specific-to-each-test-sample variance are probably not reflecting variance attributable to traits of substantive interest in most practical testing situations.

The imposement of the expected test sample covariance on the expected test sample variance leaves one with a residual which is the difference between the expected variance and the expected covariance. By adopting this additive model, one has implicitly assumed two orthogonal components, a component due to the covariance, and another component due to other less common elements, specificity and a random error component. The last component is thus a mixed-up component.

The expected random composite variance-covariance matrix.

From what was said about the inferred structure of an expected test sample variance in the previous section,it should come as no surprise that it is possible to conceive of an expected random composite variance-covariance matrix.

Insert Table 1 about here

By adopting the test theory assumptions as developed by Lord and Novick (1968),Chapter 8,it can be shown that the expected variance-covariance matrix takes the form presented in Table 1.From this table one can derive two interesting intraclass correlation coefficients.

Expected alpha(1) and the structure of the expected test sample variance.

By having imposed the expected covariance on the expected test sample variance,one is left with a structured test sample variance.As is evident from Table 1,this structure can be written,

$$E(\sigma(i)^2) = E(\sigma(i,j)) + E(\sigma(res)^2) \qquad (5)$$

From (5) one can find what proportion of the expected test sample variance is accounted for by expected common variance, or covariance,

$$E(\rho(I)) = E(alpha(1)) = \frac{E(\sigma(i,j))}{E(\sigma(i,j)) + E(\sigma(res)^2)} = \frac{E(\sigma(i,j))}{E(\sigma(i)^2)} \qquad (6)$$

Both in structure and content (6) is an intraclass correlation.Structurally,(6) is explained or systematic variance over observed variance.As to content,(6) describes to what extent observations within a class go together,on the avarage. In the present context (6) may be interpreted to describe how capable one average test sample observation is in tapping a defined universal trait.Structurally and functionally,

(6) is equivalent to alpha(1),as defined by Cronbach,Rajaratnam,and Gleser (1963).

An estimation of alpha(1) can be obtained by taking the average observed test sample covariance in a random composite over the average test sample variance from the same random composite,

$$\text{alpha}(1) = \frac{\bar{c}(i,j)}{\bar{v}(i)}, \text{where } i \neq j \qquad (7)$$

Expected alpha(k) and the structure of the expected composite variance.

More often than being concerned with alpha(1),one is interested in knowing to what extent a linear combination of k random test samples from a defined universe of observations is measuring some common construct.One is interested in estimating how much common variance is running through a random-parallel composite.

A composite of k test samples has a total test variance which is the sum of the elements in the test sample variance-covariance matrix.There are k test sample variances and k(k-1) covariances in this matrix.

A latent structure of an expected variance-covariance matrix is now conceived of,based on the previous considerations of how the expected test sample variance is structured,and shown in Table 1.The best guess one can make as to the covariances among randomly drawn test samples is the expected covariance in the universe.Thus,k^2 expected covariances are defined into the matrix.According to the rationale developed for the test sample variance structure,one should note that the expected covariance also goes into the principal diagonal.Into the principal diagonal was also defined the expected residual test sample variance as an addition to the expected covariance.

The expected composite variance can now be written as a sum of the conceptually structured variance-covariance matrix,

$$E(\sigma(x)^2) = k^2E(\sigma(i,j)) + kE(\sigma(res)^2) \tag{8}$$

From (8) one can now find the proportion of common variance running through the composite to total composite variance by taking the weighted expected covariance over the sum of the expected variance-covariance matrix, which is equal to expected test variance,

$$E(alpha(k)) = \frac{k^2E(\sigma(i,j))}{k^2E(\sigma(i,j)) + kE(\sigma(res)^2)} \tag{9}$$

Though not the traditional form of alpha, (9) is nevertheless a form of the alpha coefficient. This will become more understandable when we are going to discuss the relationship between the covariance approach and the analysis of variance approach. Functionally, (9) is common variance over total variance, here expressed as an expected value because one is committed to a random-parallel conception of composites.

In relation to (6), (9) is a stepped-up intraclass correlation, which might be represented by the symbol $\rho(k)$.

By rewriting the denominator in (9) as expected composite variance, one arrives at an alpha(k) form which is structurally the Tryon form of the reliability of an unstratified composite,

$$E(alpha(k)) = \frac{k^2E(\sigma(i,j))}{E(\sigma(x)^2)} \tag{10}$$

(10) is identical to (4) which is a more generally conceived Tryon form. However, while (4) was derived by conceiving of a correlation among two random-parallel composites, (10) is derived by conceiving of a latent variance-covariance structure of <u>one</u> expected composite variance-covariance matrix only.

In estimating (10), one would compute the average covariance amomg observed test samples in a random composite and multiply by k^2 to get the numerator. The obtained test variance of a random composite should be the proper estimate of the denominator. Thus, the estimation form of (10) becomes,

$$alpha(k) = \frac{k^2\bar{c}(i,j)}{V(x)}, \tag{11}$$

which is Tryon's general form.While (11) is an exact determination of Tryon's reliability coefficient under his more restricted random-sampling assumptions,it is in the present context an estimation form of the expected internal consistency of a random-parallel composite with full observation of the random-sampling assumptions.

The estimation form (11) is inconvenient as a computing form.For computing convenience the traditional alpha form is the best one and can easily be derived from (11),

$$\text{alpha}(k) = \frac{k^2\bar{c}(i,j)}{V(x)} = \frac{k^2\Sigma C(i,j)}{k(k-1)V(x)} = \frac{k \quad \Sigma C(i,j)}{(k-1) \quad V(x)}$$

$$= \left(\frac{k}{k-1}\right)\left(1 - \frac{\Sigma v(i)}{V(x)}\right) \qquad (12)$$

While the traditional alpha form as given by (12) is a convenient computing formula,it should though be clear that (12) is more obscure as a defining formula than is (10). The basic structural properties of alpha is reflected in (10),while (12) is a derived form where the multiplier k/(k-1) is arrived at by reducing.The obscurity of traditional alpha seems to be connected with this multiplier, which sometimes is being interpreted without making the meaning of alpha clearer.(For a recent example,see Nunnally (1967),195).

One of the most successful endeavors to make clear the meaning of the traditional alpha form might be Cronbach's comments,

> "This important relationship states a clear meaning for alpha as k/(k-1) times the ratio of interitem covariance to total covariance.The multiplier k/(k-1) allows for the proportion of variance in any item which is due to the same elements as the covariance."(Cronbach 1951,305)

According to the rationale as developed for arriving at (9),it is clear that Cronbach's comments reflect an intuition which,although not developed formally,is nevertheless correct.

Numerical example.

Given a hypothetical data matrix, Table 2, where rows designate persons and columns test samples. According to the

Insert Table 2 about here

previously developed rationale one has to compute the average covariance among test samples and the average test sample variance to estimate the expected values. Then, in order to find the residual variance one has to subtract the

Insert Table 3 about here

average covariance from the average test sample variance. From Table 3 the average covariance is found to be 1,358 and the average test sample variance 1,725. Consequently, the residual is 1,725-1,358 = 0,367.

An estimate of the structured expected test sample variance is thus,

$\bar{v}(i) = \bar{c}(i,j) + \bar{v}(res) = 1{,}358 + 0{,}367 = 1{,}725.$

While the average test sample variance is 1,725; 1,358 of the total test sample variance is considered explained variance, "due to the same elements as the covariance", to quote Cronbach once again.

The ratio of the average covariance to the average test sample variance is an intraclass correlation coefficient, or alpha(1),

$\rho(I) = \text{alpha}(1) = 1{,}358/1{,}725 = 0{,}787,$

which is the estimated capability of one average test sample to tap a common-to-a-defined-universe trait.

The sum of the variance-covariance matrix, Table 3, is 23,2, which is the variance of the sum scores, X(t), Table 2. According to the defined covariance structure as given by (8), the partitioning of the composite variance give,

$$V(x) = k^2\bar{c}(i,j) + k\bar{v}(res) = 16(1{,}358) + 4(0{,}367)$$
$$= 21{,}73 \quad + \quad 1{,}47 \quad = \quad 23{,}20$$

Common variance running through the composite, to total composite variance gives,

$$\text{alpha}(k) = \frac{k^2\bar{c}(i,j)}{V(x)} = \frac{21,73}{23,20} = 0,937$$

alpha(k) tells that 0,937 of the composite variance can be considered due to a genuine difference among persons in the common trait measured by the instrument. (Insert Table 4)

For comparison, an analysis of variance of the same data matrix is presented in Table 4. This is a Hoyt analysis. By applying the formulas as given by Winer (1962), 124-132, one obtains coefficients corresponding to alpha(1) and alpha(k),

$$\text{alpha}(1) = \frac{5,800 - 0,367}{5,800+3(0,367)} = 0,787$$

$$\text{alpha}(k) = \frac{5,800 - 0,367}{5,800} = 0,938$$

The results obtained in the analysis of variance approach are identical to what is obtained by the covariance approach.

As it is believed that it may be difficult to see how the two approaches converge, next an exploration into this convergence is appropriate.

The relationship between the covariance approach and the analysis of variance approach.

In the analysis of variance approach to alpha, the E(MS) for persons reflects the conceptual variance structure,

$$E(MS(p)) = k\sigma(p)^2 + \sigma(pi)^2 \qquad (13)$$

In the covariance approach the expected composite variance is structured,

$$E(\sigma(x)^2) = k^2E(\sigma(i,j)) + kE(\sigma(res)^2) \qquad (14)$$

There is an obvious structural similarity between (13) and (14), although coefficients differ and components are seemingly different.

The different coefficients used in the two models may be said to be a matter of convention. In applying analysis of variance techniques to test data, the person variance, symbolized as MS(p), is 1/k of the total composite variance as computed from the sum scores. Again, the variance of the average person score is 1/k of MS(p), and $1/k^2$ of V(x), the sum score variance. The formal relationship between these three ways of defining person variance written as an equality, is thus,

$V(x) = kMS(p) = k^2V(\overline{x})$ (See, for example, Winer 1962, 125)

Therefore, recalling (13) and (14),

$$1/k(E(\sigma(x)^2) = kE(\sigma(i,j)) + E(\sigma(res)^2) = k\sigma(p)^2 + \sigma(pi)^2 \quad (15)$$

Next, it should be shown that the components in the two approaches are equal. Gulliksen (1950), 54, has shown that the residual variance in the analysis of variance,

$$\sigma(pi)^2 = \overline{v}(i) - \overline{c}(i,j) \quad (16)$$

According to (16) the interaction term, $\sigma(pi)^2$, can be written as a function of the average test sample variance and the average test sample covariance. Thus, (16) is an estimate of,

$$\sigma(pi)^2 = E(\sigma(i)^2) - E(\sigma(i,j)) \quad (17)$$

Recalling that MS(p) is 1/k of total composite variance, it should be clear that MS(p) must equal 1/k of the sum of the composite variance-covariance matrix. Also, taking advantage of what Gulliksen has shown concerning the residual variance, one can rewrite the two relevant E(MS)'s of the analysis of variance table this way,

$$E(MS(p)) = 1/k(kE(\sigma(i)^2) + k(k-1)E(\sigma(i,j))) = E(\sigma(i)^2) + (k-1)E(\sigma(i,j))$$

$$E(MS(pi)) = E(\sigma(i)^2) - E(\sigma(i,j))$$

Again, recalling that $E(MS(p)) = k\sigma(p)^2 + \sigma(pi)^2$, which is the conceptual structure of person variance, one can do this,

$$k\sigma(p)^2 = (E(\sigma(i)^2) + (k-1)E(\sigma(i,j))) - (E(\sigma(i)^2) - E(\sigma(i,j)))$$
$$= E(\sigma(i)^2) + kE(\sigma(i,j)) - E(\sigma(i,j)) - E(\sigma(i)^2) + E(\sigma(i,j))$$
$$= kE(\sigma(i,j)) \quad (18)$$

From (18) it is apparent that the variance component for persons in the analysis of variance approach, $\sigma(p)^2$, is equal to the expected covariance among test samples, $E(\sigma(i,j))$.

In the analysis of variance approach the deviation score, i.e. a residualized deviation score, is partitioned into two components, a mean score component and a residual component, due to the interaction between persons and test samples, or items. In the covariance approach one operates directly on the observed deviation score

In the analysis of variance the deviation scores of the column-centered data matrix, as presented in Table 5, are

Insert Table 5 about here

considered to contain two orthogonal components for each person/test sample combination. As mentioned, one is here concerned with a residualized data matrix where only person-relevant variance is involved. The total sum of squares of the residualized matrix is now a sum of the person components squared and the residual components squared.

The MS(p) obtained on the basis of the mean observed person score is in a test theory context regarded as an inflated universe score variance because it is an observed value, containing error. Therefore, the observed person variance, MS(p), has to be corrected by another variance supposed to be an estimate of the inflation involved. In the analysis of variance model at issue in the present context, the interaction term serves this purpose.

What is here described in common sense language, compares in the formal model to the expected mean square for persons,

$$E(MS(p)) = k\sigma(p)^2 + \sigma(pi)^2,$$

where $k\sigma(p)^2$ is estimated by $MS(p) - MS(pi)$.

In the covariance approach to alpha one can establish the variance-covariance matrix from the residualized deviation score matrix in Table 5. The deviation score is the sum of the two components in Table 5.

In computing the covariances among observed test samples one obtains in one step what is obtained in a two-step procedure in the analysis of variance. The average covariance obtained by this method is an estimate of the expected covariance, which is equivalent to the test theory result that observed score covariance equals universe score covariance. A formal proof of this result, and the assumptions made, can be found in Novick and Lewis (1967), 4.

The previous proof and discussion should leave one with the impression that the socalled variance component in the analysis of variance of test data, reflecting differences in universe scores among persons, is a misnomer for what should appropriately be called a covariance component. It is believed to facilitate a clearer understanding of what is implied in the technically sophisticated analysis of variance approach to alpha by pointing to the variance component for persons as the expected covariance among universe scores.

Concluding remarks.

While coefficient alpha, i.e. observed alpha, is an underestimate of the defined reliability or generalizability coefficient, it can likewise be shown that observed alpha is an underestimate of expected alpha.

Expected alpha is equal to the reliability coefficient and to the expected correlation between two randomly drawn composites. Thus the equality that existed within classical test theory on the observed level, but not within a random-sampling test theory model, this equality is within the latter model restored on the expected level, such that

$$E(\text{alpha}) = E(\rho(x,x')) = E(\rho(X,T)^2).$$

Expected alpha as here derived from the expected variance-covariance matrix of a random composite is shown to be an internal consistency construct per se, not necessarily dependent upon conceiving of it in terms of the expected intercorrelation among random composites nor upon the expected correlation between observed score and true score, or universe score.

References

Cronbach, L.J. Coefficient alpha and the internal structure of tests. Psychometrika, 1951, 16, 297-334.

Cronbach, L.J., Rajaratnam, N., and Gleser, G.C. Theory of generalizability: A liberalization of reliability theory. British Journal of Statistical Psychology, 1963, 16, 137-163.

Edwards, A.L. A note on Tryon's measure of reliability. Psychometrika, 1959, 24, 257-260.

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

Hoyt, C. Test reliability estimated by analysis of variance. Psychometrika, 1941, 6, 153-160.

Jackson, R.W.B., and Ferguson, G.A. Studies on the Reliability of Tests. Bulletin 12, Department of Educational Research, Toronto: University of Toronto, 1941.

Kelley, T.L. The reliability coefficient. Psychometrika, 1942, 7, 75-83.

Kuder, G.F., and Richardson, M.W. The theory of the estimation of test reliability. Psychometrika, 1937, 2, 151-160.

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

Novick, M.R. and Lewis, C. Coefficient alpha and the reliability of composite measurements. Psychometrika, 1967, 32, 1-13.

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

Tryon, R.C. Reliability and behavior domain validity: Reformulation and historical critique. Psychological Bulletin, 1957, 229-249.

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

1) Most of the work on this paper was completed while the author was a Research Associate at State University of New York at Buffalo, 1968-1969.

2) Test sample and composite are used for item and test, consistent with Tryon's (1957) use of the terms.

3) For alternative alpha formulas, see Tryon (1957) and Edwards (1959).

Table 1

Expected Variance-Covariance Matrix of a Random Composite

	1	2	3	4
1	$\sigma(i,j)+\sigma(res)^2$	$\sigma(i,j)$	$\sigma(i,j)$	$\sigma(i,j)$
2	$\sigma(i,j)$	$\sigma(i,j)+\sigma(res)^2$	$\sigma(i,j)$	$\sigma(i,j)$
3	$\sigma(i,j)$	$\sigma(i,j)$	$\sigma(i,j)+\sigma(res)^2$	$\sigma(i,j)$
4	$\sigma(i,j)$	$\sigma(i,j)$	$\sigma(i,j)$	$\sigma(i,j)+\sigma(res)^2$

Table 2

Hypothetical Data Matrix

	1	2	3	4	
a	4	5	4	5	18
b	3	4	5	4	16
c	4	4	3	3	14
d	2	3	3	2	10
e	1	2	1	2	6
	14	18	16	16	64

Table 3

Sample Covariance Matrix

	1	2	3	4
1	1.70	1.40	1.30	1.30
2	1.40	1.30	1.35	1.35
3	1.30	1.35	2.20	1.45
4	1.30	1.35	1.45	1.40

Table 4

Analysis of Variance of Hypothetical Data

Source	df	SS	MS	E(MS)
Persons	4	23.2	5.800	0.367 + 4(1.358)
Test samples	3	1.6	0.533	
P x T	12	4.4	0.367	0.367
Total	19	29.2		

Table 5

Decomposed Residualized Deviation Scores

	1	2	3	4
a	1.3 - 0.1	1.3 + 0.1	1.3 - 0.5	1.3 + 0.5
b	0.8 - 0.6	0.8 - 0.4	0.8 + 1.0	0.8 + 0.0
c	0.3 + 0.9	0.3 + 0.1	0.3 - 0.5	0.3 - 0.5
d	-0.7 - 0.1	-0.7 + 0.1	-0.7 + 0.5	-0.7 - 0.5
e	-1.7 - 0.7	-1.7 + 0.1	-1.7 - 0.5	-1.7 + 0.5