Psychological Bulletin
1988, Vol. 103, No. 2, 276-279

0033-2909/88/$00.75

The Eigenvalues-Greater-Than-One Rule and the Reliability of Components

Norman Cliff
University of Southern California

A commonly used criterion for the number of factors to rotate is the eigenvalues-greater-than-one rule proposed by Kaiser (1960). It states that there are as many reliable factors as there are eigenvalues greater than one. The reasoning is that an eigenvalue less than one implies that the scores on the component would have negative reliability. I show here that this rule is the result of a misapplication of the formula for internal consistency reliability. I also present a formula for the reliability of a component; it depends on the eigenvalue and the reliability of the individual measures.

One of the most difficult decisions that a factor analyst must make is the number of factors to retain. For the person using factor analysis in an exploratory way and using the principal-components and principal-factors methods that have been the mainstay since the late 1950s, this decision is typically made on the basis of the size and relative size of the eigenvalues. It is this methodology that is discussed here, in spite of the fact that it is now being replaced by quasi-inferential methods, such as maximum likelihood.

A simple rule, one that has been the default option of many statistical packages, is to retain as many factors as there are eigenvalues greater than unity in a principal-components analysis. If a common-factors analysis based on estimated communalities is preferred, then a two-stage analysis is necessary: a principal-component analysis to find the number of factors, followed by a principal-factors analysis to compute the loadings.

There are three rationales for the rule. One is due to Guttman (1954), who proved that *in the population* it provided a lower bound for the number of common factors. The second is the informal reasoning that a component is of little interest if it accounts for less variance than a single variable does. The third is the statement by Kaiser (1960) that a component score will have negative reliability if the eigenvalue is less than unity. The purpose of the present article is to show that the last rationale was based on an erroneous application of a commonly used formula for estimating the reliability of a multi-item test.

Zwick and Velicer (1986) recently reviewed the literature on rules for identifying the "correct" number of factors and presented additional Monte Carlo data of their own. They concluded that the Kaiser rule tends to overestimate the number of factors in sample correlation matrices, whereas Guttman (1954) showed that it is a lower bound, that is, tends to underestimate the number of factors. Guttman's conclusion, however, was developed with respect to the population correlation matrix, whereas most data decisions are based on sample data, and the Monte Carlo studies cited by Zwick and Velicer simulate sample correlation matrices. It is in fact relatively simple to construct population matrices in which the rule will underestimate the number of factors; Table 1 is an example, as will be seen in more detail later. It is nearly as simple to construct population correlation matrices whose sample estimates will lead to overestimates of the number of factors if the rule is applied. Thus, the rule can systematically either over- or underestimate the number of factors, depending on circumstances.

Such negative evidence leads to questioning the basis for the rule. In what follows, I show that Kaiser's (1960) rationale for relating the reliability of components to the number of eigenvalues greater than unity was based on a misapplication of a common formula for the reliability of a composite and that the reliability of a principal component depends on the reliability of the measures. Thus, reliability of components cannot be deduced from the size of the eigenvalues, and the Kaiser rationale for retaining as many components or factors as there are eigenvalues greater than unity does not have any logical basis. This article focuses on only that issue and does not deal further with the other two justifications for the eigenvalues-greater-than-one rule.

Kaiser (1960) Rationale

The suggestion by Kaiser (1960) was based on a commonly used formula for the reliability of a total score. Kuder and Richardson (1937) first suggested the formula, and it is called *Kuder-Richardson formula 20* (K-R 20). It and its rationale were elaborated by Cronbach (1951), who called the generalization *coefficient alpha.*

The formula estimates the parallel form reliability of a total score from the internal consistency among subscores or items. It is extremely useful because it does not require the actual computation of correlations among the items. A common formulation is

$$r_{xx'} = p/(p-1)\,\{(s_x^2 - \sum_j s_j^2)/s_x^2)\}. \qquad (1)$$

Correspondence concerning this article should be addressed to Norman Cliff, Psychology Department, University of Southern California, Los Angeles, California 90089-1061.

That is, the estimated correlation between a total score x and a hypothetical parallel total score x' depends on just p (the number of items), the variance s_x^2 of x, and the sum of the variances of the items $\sum_j s_j^2$.

Kaiser (1960) proposed that this formula was applicable to deciding on the number of factors. If a factor is to be retained, the corresponding composite score should be reliable. At least, it should have positive reliability in the internal consistency sense.

His rationale seems to have been as follows:[1] The cth principal component is a weighted composite score in which the weights are the elements $w_{1c}, w_{2c}, \ldots, w_{jc}, \ldots, w_{pc}$ of the cth eigenvector, and the variance of the composite is the corresponding eigenvalue. The term $\sum_j s_j^2$ in Equation 1 actually represents $\Sigma\ w_{jc}^2 s_j^2$, but in K-R 20 all the weights are equal to unity; hence, it is simplified to $\sum_j s^2$.

In the factor-analytic context, the weights are unequal, but a simplification still occurs because there the s_j^2s represent the variables' variances, and here these are all equal to 1.0. Also, eigenvectors are conventionally normalized so that

$$\sum_j w_{jc}^2 = 1.0.$$

Therefore, the reliability of the cth principal component can, by substitution, supposedly be estimated as

$$r_{cc'} = \frac{p}{p-1}\,\frac{\lambda_c - \sum w_{jc}^2(1.0)}{\lambda_c} = \frac{p}{p-1}\,\frac{\lambda_c - 1.0}{\lambda_c}. \quad (2)$$

If $\lambda_c < 1.0$, the estimated reliability is negative, so the component is supposedly unreliable. Therefore, one should retain only as many factors as there are principal-components eigenvalues greater than unity; hence, the rule. I show below that there is a flaw in this application of K-R 20.

Reliability of Composite Scores

Equation 1 is a very special case of a general formula for the correlation between two composites. The general formula for the correlation between one composite score, $x_c = \sum_j w_{cj} y_j$, and another, $x_{c'} = \sum_j w_{c'j} y_j$, is by definition

$$r_{cc'} = \frac{\sum_j \sum_k w_{cj}\, w_{c'k} s_{jk}}{(\sum_j \sum_k w_{cj}\, w_{ck} s_{jk})^{1/2} (\sum_j \sum_k w_{c'j}\, w_{c'k} s_{j'k})^{1/2}}. \quad (3)$$

The formula gives the correlation between one combination, x_c, and a second combination, $x_{c'}$. There is a vector of weights for the c combination—$w_{c1}, w_{c2}, \ldots, w_{cj}, \ldots, w_{cp}$—and one for the c'—$w_{c'1}, w_{c'2}, \ldots, w_{c'j}, \ldots, w_{c'p}$.

In the case of computing the reliability of a composite, there are two parallel batteries with p measures in each, and the same weights are applied to each battery. The reliability is the correlation between a weighted composite of one battery and the same weighted composite of the other. Then a matrix version of Equation 3 becomes

$$r_{cc'} = \frac{\mathbf{w}'\mathbf{S}_{12}\mathbf{w}}{(\mathbf{w}'\mathbf{S}_{11}\mathbf{w})^{1/2}(\mathbf{w}'\mathbf{S}_{22}\mathbf{w})^{1/2}}, \quad (4)$$

where $\mathbf{S}_{12}$ is the matrix of covariances between measures in the first battery and measures in the second, $\mathbf{S}_{11}$ is the matrix of covariances among members of the first battery, $\mathbf{S}_{22}$ is the matrix of covariances among members of the second, and $\mathbf{w}$ is the vector of weights.

Equation 1 is a special case of Equation 4 provided certain simplifications hold: (a) All the weights are equal; that is, x is a simple total score; (b) the variances of x and x' are the same; (c) the average covariance between a measure in one battery and a measure in the other battery is the same as the average covariance among measures in the same battery. Gulliksen (1950), for example, showed the steps in deriving Equation 1 from Equation 4 under these assumptions. The equation then gives the correlation between the total score on the available battery of items or measures and the total score on a hypothetical second one and, as such, estimates a parallel form reliability.

Equation 4 does not reduce to Equation 1, however, unless the assumptions in the preceding paragraph hold. But in applying Equation 2, the weights are the elements of an eigenvector that are typically unequal, so the conclusion made by Kaiser (1960) is erroneous: There is no direct relation between the size of an eigenvalue and the reliability of the corresponding composite.

Reliability of Principal Components

A much weaker version of the rule can be derived, but it requires stronger assumptions about the covariances between variables and requires estimates of the individual measures' reliabilities. When the weights are unequal, one can only estimate the reliability of the composite by assuming that the second battery of measures is parallel measure for measure with the first. One can assume that if $y_{j'}$ and $y_{k'}$ are the corresponding parallel measures in the second battery (which is hypothetical in this application) that are respectively parallel to the observed y_j and y_k, then $s_{jk} = s_{j'k'} = s_{jk'} = s_{j'k}$, and $s_j^2 = s_{j'}^2$. That is, the covariances between corresponding measures are the same whether both are in one battery or one is in one battery and one is in the other. This assumption results in the two standard deviation terms in the denominator of Equation 4 being equal. (It is possible to make a somewhat weaker assumption, but it involves the weights, is very complicated to state, and seems very unrealistic in practice.)

The correlation between the observed component and the hypothetical second component is not yet defined because, although it is assumed that $s_{j'k}$ (the covariance between the ob-

[1] Kaiser (1960) did not specify the formula or formulas that he used, stating, "I have worked out all the formulas for Kuder-Richardson reliability of factors" (p. 145).

served jth item and the hypothetical kth item) is equal to the observed s_{jk}, there is no assumption about the covariance between corresponding measures, $s_{jj'}$, that is, about the reliability of the individual measures.

In the case of K-R 20 or alpha, the average covariance between a measure in one battery and its parallel measure in the other is assumed to be equal to the average covariance within the battery. This leads to Equation 1. In the case of unequal weights, the results are different. To see what happens, one can express Equation 4 in terms of the observed covariances. It is also useful to separate out the terms involving corresponding variables in the numerator.

$$r_{cc'} = \frac{\sum_j \sum_{k \neq j} w_{cj} w_{ck} s_{jk} + \sum_j w_{cj}^2 s_{jj'}}{\sum_j \sum_k w_{cj} w_{ck} s_{jk}}. \quad (5)$$

The quantity $\Sigma w_j^2 s_j^2$ can be added and subtracted in the numerator:

$$r_{cc'} = \frac{(\sum_j \sum_{k \neq j} w_{cj} w_{ck} s_{jk} + \sum_j w_{cj}^2 s_{jj'}) - \sum w_{cj}^2 s_j^2 + \sum w_{cj}^2 s_{jj'}}{\sum \sum w_{cj} w_{ck} s_{jk}}. \quad (6)$$

The part of the numerator in parentheses is the same as the denominator. Also, if the variables are measured in standard scores, one can substitute the correlations $r_{jj'}$ for the covariances $s_{jj'}$ and 1.0 for s_j^2. When w_c is an eigenvector, two additional things happen: The denominator becomes the eigenvalue λ_c, and $\Sigma w_{cj}^2 s_j^2 = 1.0$. Therefore,

$$r_{cc'} = \frac{\lambda_c - 1 + \sum w_{cj}^2 r_{jj'}}{\lambda_c}. \quad (7)$$

This is the crucial difference between Kaiser's (1960) suggestion and the present analysis. The correct numerator is not $\lambda_c - 1$ but $\lambda_c - 1$ plus the weighted sum of the reliabilities. This is where the eigenvalues-greater-than-one rule goes astray.

To get some feel for $r_{cc'}$ or to use it in applied work, Σw_{cj}^2 can be reinserted for the 1.0; then a little manipulation shows that

$$r_{cc'} = \frac{\lambda_c - \sum w_{cj}^2 (1 - r_{jj'})}{\lambda_c}. \quad (8)$$

That is, the reliability will be positive if λ_c is greater than the weighted sum of the error variances. If the reliabilities for different variables are equal, then Equation 8 simplifies to

$$r_{cc'} = \frac{\lambda_c - (1 - r_{jj'})}{\lambda_c}. \quad (9)$$

Thus, the composite reliability would be negative only for eigenvalues less than the complement of the reliability of a single measure.

Table 1
Hypothetical Correlation Matrix

Variable	1	2	3	4	5	6
1	—	.80	.80	.60	.60	.60
2		—	.80	.60	.60	.60
3			—	.60	.60	.60
4				—	.80	.80
5					—	.80
6						—

Example

The conclusions presented here can be illustrated with a simple example. Table 1 contains the intercorrelations among six variables. They clearly fall into two groups, supporting the conclusion that there should be two factors. The largest eigenvalue is 4.4, but the second is only 0.8, so there is only one factor by the Kaiser-Guttman rule, contradicting the clear impression provided by the correlations.

One can see that the rule does not say how many components result in reliable component scores by assuming various values for the reliability of the individual variables. If one makes the conservative assumption that each reliability is equal to .68, which is the average correlation between the variables, then Equation 5, 8, or 9 will show that the reliability of the second composite is .60, not −.30 as the application of the K-R 20 formula would imply. More realistic reliabilities of .80 and .88 lead to composite reliabilities of .75 and .85, respectively, and only when the variables' reliabilities reach an unrealistically low .20 does the composite reliability reduce to zero. Thus, the reliability of a composite will often be positive if an eigenvalue is less than 1.0 and, in any event, depends on the reliability of the individual measures.

Discussion

Thus, comparing the eigenvalues to unity in a principal-components analysis does not show how many components will be reliable. Perhaps it is surprising that the relation between reliability and composites was never questioned on commonsense grounds. Consider the case of a battery of perfectly reliable but completely uncorrelated measures. All the eigenvalues will be equal to unity; that is, the components will have zero reliability by the rule, yet all the possible composites of the measures will be perfectly reliable.

The present conclusion is that there will be more reliable components than indicated by the eigenvalues-greater-than-one rule, and the small example suggests that there can be more common factors than the rule indicates. On the other hand, Zwick and Velicer (1986) found that the majority of the Monte Carlo studies, including their own, show that the rule overestimates the number of common factors.

One can resolve the seeming conflict by reference to several distinctions. First, as already noted, the number of reliable components and the number of common factors are different

questions, questions that perhaps never should have been confused. Second, the Guttman (1954) rationale for the number of common factors, which was rigorously proved, applies to the population. Sampling effects tend to increase the number of eigenvalues greater than one, so the rule tends to overestimate the number of factors when applied to sample matrices.

I hope the present article will contribute to the discontinuance of the simple acceptance of the eigenvalues-greater-than-one rule as defining the number of factors.

References

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297–334.

Gulliksen, H. (1950). *Theory of mental tests.* New York: Wiley.

Guttman, L. (1954). Some necessary and sufficient conditions for common factor analysis. *Psychometrika, 19,* 149–161.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20,* 141–151.

Kuder, G. F., & Richardson, M. W. (1937). The theory of estimation of test reliability. *Psychometrika, 2,* 151–160.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99,* 432–442.

Received June 23, 1986
Revision received May 4, 1987
Accepted May 11, 1987 ■

Women and Ethnic Minority Men and Women: Reviewers for Journal Manuscripts Wanted

If you are interested in reviewing manuscripts for APA journals, the APA Publications and Communications Board would like to invite your participation. Manuscript reviewers are vital to the publication process. As a reviewer, you would gain valuable experience in publishing. The P&C Board is particularly interested in encouraging women and ethnic minority men and women to participate more in this process.

If you are interested in reviewing manuscripts, please write to Leslie Cameron at the address below. Please note the following important points:

- To be selected as a reviewer, you must have published articles in peer-reviewed journals. The experience of publication provides a reviewer with the basis for preparing a thorough, objective evaluative review.
- To select the appropriate reviewers for each manuscript, the editor needs detailed information. Please include with your letter your vita. In your letter, please identify which APA journal you are interested in and describe your area of expertise. Be as specific as possible. For example, "social psychology" is not sufficient—you would need to specify "social cognition" or "attitude change" as well.
- Reviewing a manuscript takes time. If you are selected to review a manuscript, be prepared to invest the necessary time to evaluate the manuscript thoroughly.

Write to Leslie Cameron, Journals Office, APA, 1400 N. Uhle Street, Arlington, Virginia 22201.