Multivariate Behavioral Research

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hmbr20

Very Simple Structure: An Alternative Procedure For Estimating The Optimal Number Of Interpretable Factors

William Revelle & Thomas Rocklin

Version of record first published: 10 Jun 2010.

Multivariate Behavioral Research, 1979, 14, 403-414

VERY SIMPLE STRUCTURE: AN ALTERNATIVE PROCEDURE FOR ESTIMATING THE OPTIMAL NUMBER OF INTERPRETABLE FACTORS

WILLIAM REVELLE and THOMAS ROCKLIN
Northwestern University

ABSTRACT

A new procedure for determining the optimal number of interpretable factors to extract from a correlation matrix is introduced and compared to more conventional procedures. The new method evaluates the magnitude of the Very Simple Structure index of goodness of fit for factor solutions of increasing rank. The number of factors which maximizes the VSS criterion is taken as being the optimal number of factors to extract. Thirty-two artificial and two real data sets are used in order to compare this procedure with such methods as maximum likelihood, the eigenvalue greater than 1.0 rule, and comparison of the observed eigenvalues with those expected from random data.

A frequent point of concern in measurement is the proper number of constructs to measure. Given a particular domain of items or of tests, what is the best way to describe the domain? Is it better to have a few, broad factors, or many, narrow ones? This trade off between parsimony and completeness, or between simplicity and complexity is debated frequently. In psychometrics we call this the number of factors problem. As would be expected, there is a wide variety of proposed solutions to this problem which may be grouped into three major approaches: the use of theoretical arguments, psychometric rules of thumb, and statistical estimates of goodness of fit. We would like to introduce a procedure which is a conglomerate of all three approaches and to compare this procedure to a variety of other decision rules for determining the number of factors to extract from a given data set. To do this we first will outline our procedure and then make comparisons on thirty-two artificial data sets with known structure and two real data sets with inferred structures.

As we have already noted, there are three major ways to determine how many factors to extract to summarize and describe a particular data set. Of these three, theoretical principles are probably the simplest; extract and rotate factors only as long as

This research was supported in part by Grant #MH 29209-02 from the National Institute of Mental Health. Computing funds for this project were generously provided by the Office of Research and Sponsored Programs, Northwestern University. An early version of this paper was presented at the European meeting of the Psychometric Society in Uppsala, Sweden, 1978. Reprint requests should be sent to William Revelle, Department of Psychology, Northwestern University, Evanston, Illinois 60201.

they are interpretable. Such a procedure may be operationalized as listing the salient items for each factor, interpreting the factor in terms of these salient loadings, and then comparing the interpretability of a k factor solution with a k−1 and a k+1 factor solution. Alternatively, theory can lead one to predict a certain number of factors; this number is then extracted and the factors are then interpreted according to the theory. Unfortunately, this scheme leads to differences in the number of factors extracted more as a function of the complexity of the factor analyst than that of the data that are factor analyzed. Moreover, how many researchers have diligently interpreted a particular factor solution only to discover that the variables were mislabelled, or that the wrong data set had been analyzed?

The second way to determine how many factors to extract is to use one of many psychometric rules of thumb. Thus, we can plot the successive magnitude of the eigenvalues and try to apply the scree test (Cattell, 1966), or we can compare the size of the eigenvalues to the expected values given random data (Montanelli and Humphreys, 1976), or we can extract as many factors as have eigenvalues greater than 1.0 (Kaiser, 1960). No one seems to agree which rule of thumb is the best, and all agree that there exist particular data sets for which each rule will fail.

The third way to determine the number of factors to extract is to pose the question in a statistical fashion; how many factors are statistically necessary to describe a particular data set? These statistical procedures are fine if we are willing to settle for a parsimonious description of the data, but they are not particularly useful for generating interpretable factors. If the issue were solely one of parsimony, then it would be very appropriate to extract factors until the resulting residual matrix did not differ significantly from a random matrix. But most users of factor analysis are not interested so much in simple data description as they are in data *interpretation.* If we were interested in merely summarizing data, then we would never bother to rotate to various criteria of simple structure. The purpose behind such rotations is to try to arrive at solutions which are theoretically useful. By this, we mean ones that allow us to interpret the factors.

But how do we define a factor solution that is interpretable? A helpful but not necessary condition for interpretability is that the solution have a simple structure. By which we mean that each factor should have some clear cut meaning in terms of its pattern of large and small factor loadings, and in addition each item should have a non-zero loading on one and only one factor.[1] Such a simple structure is particularly important if we are courageous enough (or foolish enough) to attempt to factor *item* inter-correlation matrices. When factoring tests, it is quite reasonable to assume that no test is truly unifactorial. But when factoring items, it is very helpful (and for well-written items, justifiable) to assume that an item can be accounted for by one and only one factor. This is particularly appropriate when factoring items in order to form scales in that it avoids the problem of forming scales with overlapping items.

In reality, however, the solutions of factoring and rotation are not quite what we like. That is, although most items have at least one large loading, the remaining loadings are rarely exactly zero. What happens then when we distort the factor structure matrix to look the way we like to talk about it? That is, what happens to the quality of a particular factor solution when we degrade it to the simple structure which we think is there?

To answer this question we use a procedure we call the Very Simple Structure criterion (Revelle, Note 1). As its name implies, this method *is* very simple but it has several interesting properties. The first is that it combines the question of how many factors to extract with the question of how to rotate the factors which have been extracted. The second is that it tests the hypothesis that the data truly are simple structured, or more accurately, it gives an index of how badly the data depart from simple structure.

The steps in finding the Very Simple Structure criterion are as follows:

1) Find an initial factor solution with k factors. This factor solution may be a maximum likelihood, principal factor, centroid, group factor, or any other preferred extraction procedure.

2) Rotate the solution to maximize the rotational criterion that is preferred. Such transformations include Varimax, Quartimax, or any of a variety of oblique transformations. Call this rotated factor pattern matrix $\mathbf{F}_k$.

3) Apply the Very Simple Structure criterion. Specifically:

a) For a Very Simple Structure solution of factor complexity v, replace the $k-v$ smallest elements in each row of the factor pat-

1. Note that this definition of simple structure does not perfectly agree with that proposed by Thurstone (1947) nor with Cattell's (1973) hyperplane criterion of simple structure.

tern matrix with zeros. Call this simplified factor matrix $\mathbf{S}_{vk}$. This is what we do in practice when we attempt to interpret factors by their highest loadings. It is important to note that, unless $\mathbf{F}_k$ has a simple structure of complexity v, then $\mathbf{S}_{vk}$ is not equivalent to the initial factor solution $\mathbf{F}_k$ but is a simplified form of it (Some people prefer to call $\mathbf{S}_{vk}$ a degraded form of $\mathbf{F}_k$).

b) To evaluate how well a particular rotated factor solution $\mathbf{F}_k$ fits a simple structure model of factor complexity v, consider how well the matrix:

$$[1] \qquad \mathbf{R}_v{}^* = \mathbf{S}_{vk}\Phi\mathbf{S}'_{vk}$$

(where Φ is the factor inter-correlation matrix) reproduces the initial correlations in $\mathbf{R}$. That is, find the residual matrix:

$$[2] \qquad \overline{\mathbf{R}}_v = \mathbf{R} - \mathbf{R}_v{}^* = \mathbf{R} - \mathbf{S}_{vk}\Phi\mathbf{S}'_{vk} \quad .$$

c) as an index of fit of $\overline{\mathbf{R}}_v$ to $\mathbf{R}$, find one minus the ratio of the mean square residual correlation to the mean square original correlation

$$[3] \qquad \mathrm{VSS}_{vk} = 1 - \mathrm{MS}_{\bar{r}}/\mathrm{MS}_r$$

where the degrees of freedom for these mean squares are the number of correlations estimated less the number of free parameters in $\mathbf{S}_{vk}$. The mean squares are found for the lower off-diagonal elements in $\mathbf{R}$ and $\overline{\mathbf{R}}$.

4) Finally, to determine the appropriate number of factors to extract, find the value of the Very Simple Structure criterion for all values of k from one to the rank of the matrix. The optimal number of interpretable factors (of complexity v) is the number of factors, k, which maximizes VSS_{vk}. If a very simple structure of factor complexity one is believed appropriate, then only VSS_{1k} needs to be evaluated, but it is straightforward to evaluate the entire family of simple structures. Thus, to determine the optimal number of interpretable factors to extract from a correlation matrix, it is necessary to compare the goodness of fit of the re-

duced (simple) structure matrix to the initial correlation matrix for a variety of number of factors. If the correlation matrix has a simple structure of rank k and of complexity v, then the goodness of fit of Very Simple Structure will be maximized at that value.[2]

What exactly is the Very Simple Structure test doing and why should it achieve a maximum value at the appropriate number of factors? It is degrading the initial rotated factor solution by *assuming* that the nonsalient loadings are zero, even though in actuality they rarely are. What VSS does is test how well the factor matrix we *think about* and talk about actually fits the correlation matrix. It is not a confirmatory procedure for testing the significance of a particular loading, but rather it is an exploratory procedure for testing the relative utility of *interpreting* the correlation matrix in terms of a family of increasingly more complex factor models.

The simplest model tested by VSS is that each item is of complexity one, and that all items are embedded in a more complex factor matrix of rank k. This is the model most appropriate for scale construction and is the one we use most frequently when we talk about factor solutions. More complicated models may also be evaluated by VSS. Such models allow each item to be of complexity two, three, etc. but assume that the overall matrix is of higher rank. An example of such a higher order model could be a Bi-Factor model, or an overlapping cluster model. While normally the models of a higher order will agree with the optimal number of factors identified by the order 1 model, this is not always the case. As we have shown elsewhere (Revelle, Note 1) the optimal complexity 1 solution to the Holzinger-Harman problem is different from the optimal complexity 2 solution.

How does the Very Simple Structure criterion relate to the number of factors problem? By comparing the values of VSS_{vk} for increasing values of k and for fixed v, the fit will become better as long as the correlation matrix has a simple structure of a higher rank. For example, consider a test in which the items form three independent clusters with high correlations within

2. A short FORTRAN IV program to do these analyses is available from the authors. Alternatively, it should be noted that it is possible to do conceptually similar analyses by using a combination of the EFAP and COFAMM programs of Sörbom and Jöreskog (1976). The COFAMM technique involves an exploratory factoring followed by iteratively rotating to targets with progressively fewer zero elements. This process is then terminated when tests of goodness of fit indicate that no more loadings differ significantly from zero.

these clusters but zero correlations between clusters. Clearly a 3 factor solution will be better than a 2 factor solution, but why should a 4 factor solution be worse than a 3? By assigning some items to the fourth factor, we are saying that they do not correlate with the items which have their highest loadings on one of the first three factors. But, given a 3 cluster simple structure, this is incorrect, and the size of the residuals will increase over that observed in the 3 factor case. This will result in an increase of VSS_{1k} for $k = 1$ to 3, but in a decrease for all k greater than 3.

It is important to note that to determine the number of factors, values of k should be varied for a fixed value of v. $VSS_{v+1,k}$ will always be higher than VSS_{vk}, since the latter is based on a more severe degradation of the factor solution.

Examples

Simulations

How does VSS compare to other procedures for estimating the number of factors? Although it is possible to make many comparisons, we will limit ourselves to the type of data most often found when factoring personality or ability inventories. Typically, the communalities are low, and a simple structure model is thought to be appropriate. We have considered thirty-two 24 item tests made up of items with communalities of .3. For samples of size 50, 100, 200, and 400 we have generated one, two, three and four factor structures. Two replications were generated for each combination of sample size and factor structure. Thus, in the four factor case, each factor had six salient items with loadings in the population of .55 and the remaining 18 loadings with population values of 0.0. Each of these thirty-two data sets was factored using both maximum likelihood and the principal factor extraction algorithms, and then rotated to conventional simple structure using a Varimax algorithm. For each data set, we found the number of factors by VSS, the eigenvalue greater than 1.0 rule, the Montanelli and Humphreys rule, and two varieties of maximum likelihood procedures. The first maximum likelihood estimate was simply whether or not the χ^2 was significant for that number of factors. If it was, we extracted one more, and tested again. The second was whether the addition of one more factor resulted in a significant decrease in χ^2. If it did, that additional factor was extracted.

Since the VSS criterion is claimed to achieve a maximum value at the optimal number of factors, it is useful to see what the VSS_{1k} values are for the various problems. Figure 1 shows complexity 1 solutions for these various sample sizes. Except for one problem with 3 factors and 50 subjects and one with 4 factors and 50

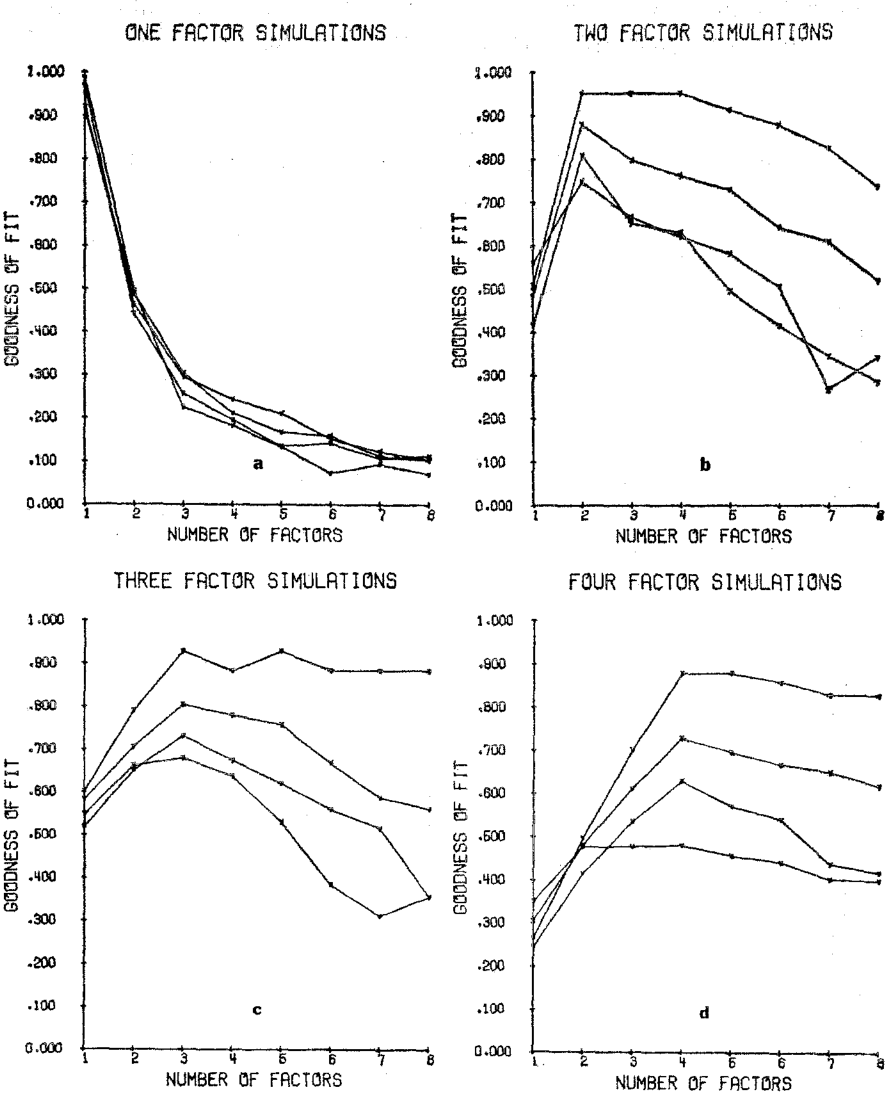

Fig. 1. Goodness of fit as a function of sample size, number of factors, and simulated factor structure. Each data point represents the mean of two replications of the same sample size and factor structure. In each case the goodness of fit of the appropriate number of factors increases across sample size ($N = 50, 100, 200, 400$).

subjects, the VSS criterion achieved its maximum value at the correct number of factors.

How well does VSS compare to the other, more established procedures? Table 1 lists for each simulation the number of factors identified by both maximum likelihood rules, by the eigenvalue greater than 1.0 rule, the Montanelli and Humphreys rule, and by Very Simple Structure. It is clear that on these simulated data sets that VSS does quite well (identifying the correct number 30

Table 1
Number of Factors Suggested by Various Methods: Simulations

Sample Size	$\lambda > 1.0$[a]	Montanelli and Humphreys	Maximim Likelihood[b]	Maximum Likelihood[c]	VSS
		One Factor Simulations			
50	7	1	1	2	1
50	8	1	1	$k < 6$	1
100	7	1	1	5	1
100	8	1	2	3	1
200	5	$k > 8$	1	6	1
200	4	1	1	5	1
400	3	1	1	N/A	1
400	3	$k > 8$	1	$4 < k < 8$	1
		Two Factor Simulations			
50	7	3	2	4	2
50	8	2	2	$k > 3$	2
100	9	3	2	$3 \leq k < 6$	2
100	7	3	4	$k \geq 8$	2
200	7	$k > 8$	2	$k > 8$	2
200	6	$k > 8$	2	$5 \leq k < 8$	2
400	3	$k > 8$	2	$4 < k < 6$	2
400	4	$k > 8$	2	5	2
		Three Factor Simulations			
50	7	4	2	$5 < k < 7$	3
50	8	5	3	5	2
100	8	3	2	3	3
100	8	$k > 8$	4	7	3
200	8	$k > 8$	3	$k > 8$	3
200	7	$k > 8$	3	$5 \leq k < 8$	3
400	4	$k > 8$	3	5	3
400	4	$k > 8$	3	8	3
		Four Factor Simulations			
50	9	7	2	6	2
50	8	5	4	$5 < k < 8$	4
100	8	$k > 8$	4	7	4
100	8	$k > 8$	6	$k > 7$	4
200	9	$k > 8$	4	6	4
200	7	$k > 8$	4	$k > 8$	4
400	5	$k > 8$	4	5	4
400	4	$k > 8$	4	6	4

Note: Ranges are given where maximum likelihood factor analysis failed to converge within a reasonable number of iterations or when the number of factors was greater than 8 for the Montanelli and Humphreys rule.

[a] λ is the size of the eigenvalue of the principal component.

[b] Test for significance of residuals.

[c] Test for significance of change in χ^2.

out of 32 times). In addition, it is also clear that for these low communality items the eigenvalue greater than 1.0 rule does very badly (1 out of 32). Conventional maximum likelihood does fairly well, properly identifying the correct number of factors 25 out of 32 times. For each method except VSS errors were in the direction of *overfactoring*. VSS erred only twice. But each time it *underfactored*, an error which some people consider more serious.

Established Scales

Finally, when introducing any new psychometric method, it is important to demonstrate that it gives reasonable results on real data as well as on artificial problems. It always is easy to cook up simulated data sets which show how well a procedure will work, but it is important to show what happens when the procedure is applied to real data problems.

The first problem we have chosen is an analysis of the factor structure of the Alpert and Haber *Debilitative and Facilitative Anxiety Scales*. The second applied data set is an analysis of the factor structure of the Eysenck Introversion/Extraversion scale from the *Eysenck Personality Inventory*.

Alpert and Haber (1960) have claimed that their scales assess two different components of anxiety. It is claimed that one factor measures levels of facilitative test anxiety, while the other factor measures levels of debilitative test anxiety. There is considerable disagreement between the number of factors indicated by maximum likelihood estimates, the eigenvalue greater than 1.0 rule, and the Very Simple Structure criterion (Table 2). VSS indicated that one factor (general test anxiety) was most appropriate. This is seen graphically in Figure 2a. In order to allow for a comparison, we have included our simulation of a one factor test in Figure 2b.[3]

The second demonstration of Very Simple Structure is the Eysenck Introversion/Extraversion scale from the EPI. Eysenck (1977) has claimed that although there are two sub-factors in the scale, it is more fruitful to consider it as a one factor test than as a two factor test. Once again, there is considerable disagreement between the number of factors indicated by the various procedures

3. VSS, of course, has two parameters: The number of factors (k) and the complexity or number of non-zero loadings for each item (v). These can be varied independently (within the constraint $v \leq k$) allowing one to evaluate a *family* of solutions. In Figure 2, for example, k was varied from 1 to 8, and v varied from 1 to k. Each curve represents a different value of k.

(Table 2). VSS indicated that two factors were most interpretable in terms of a simple structure model (Figure 2c). For comparison we have included our simulation of a two factor test in Figure 2d. In terms of the item content, these two factors represent sociability and impulsivity. As an experimental validation of a two factor solution to the Introversion/Extraversion scale, we have recently shown (Revelle, Humphreys, Simon & Gilliland, in press) that impulsivity and sociability have very different patterns of correlations with such experimental variables as caffeine-induced stress or the time of day. This experimental independence makes us much more confident that we have correctly rejected the single factor hypothesis for this scale.

Conclusions

Very Simple Structure is a procedure which combines parts of three major approaches to the number of factors problem. It makes use of theoretical arguments for simple structure, but attempts to see how well such a model actually fits the data. Rather than evaluating this fit in terms of statistical significance (although this is, of course, possible), we prefer to plot the goodness of fit value as a function of the number of factors. We believe that the optimal number of interpretable factors is the number which maximizes the Very Simple Structure criterion.

Table 2
Number of Factors Suggested by Various Methods: Real Data

Inventory	$\lambda > 1.0$[a]	Montanelli and Humphreys	Maximum Likelihood[b]	Maximum Likelihood[c]	VSS
AAT	5	$k > 8$	6	$k > 6$	1
EPI Extraversion	9	$k > 8$	3	$k > 6$	2

Note: Ranges are given where maximum likelihood factor analysis failed to converge within a reasonable number of iterations or when the number of factors was greater than 8 for the Montanelli and Humphreys rule.

[a] λ is the size of the eigenvalue of the principal component.

[b] Test for significance of residuals.

[c] Test for significance of change in χ^2.

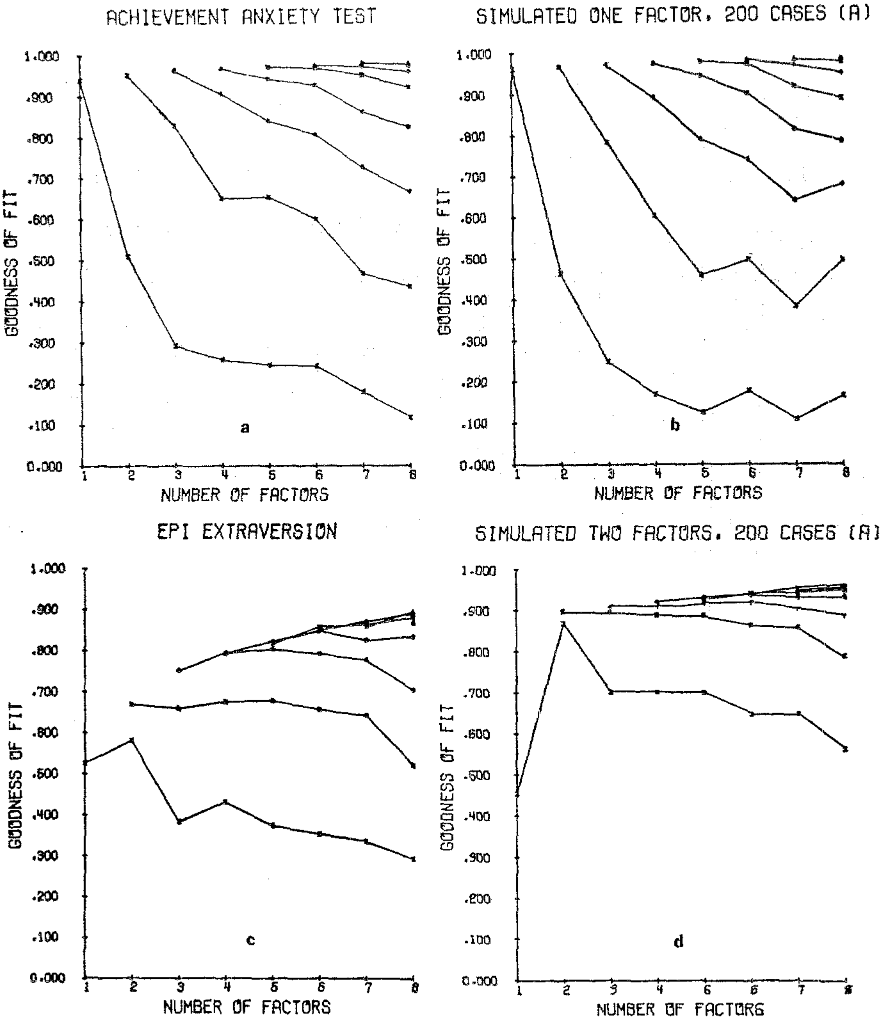

Fig. 2. Goodness of fit as a function of number of factors and factor complexity of solution. The left hand panels represent real data sets (Alpert-Haber Achievement Anxiety Test (a), and the Eysenck Personality Inventory Extraversion Scale (c)) while the right hand panels are solutions of 1 factor (b) and 2 factor (d) simulations. Each panel shows a complete VSS solution for all values of $1 \leq k \leq 8$ and $v \leq k$. The complexity (v) of the solution represented by a given line is equal to the number of factors at which that line starts.

Reference Note

Revelle, W. Very Simple Structure: An alternative criterion for factor analysis. Paper presented at the annual meeting of the Society for Multivariate Experimental Psychology, November, 1977.

REFERENCES

Alpert, R., & Haber, R. N. Anxiety in academic achievement situations. *Journal of Abnormal and Social Psychology,* 1960. *61,* 207-215.

Cattell, R. B. The scree test for the number of factors. *Multivariate Behavioral Research,* 1966, *1,* 245-276.

Cattell, R. B. *Personality and mood by questionnaire.* San Francisco: Jossey-Bass Publications, 1973.

Eysenck, H. J. Personality and factor analysis: A reply to Guilford. *Psychological Bulletin,* 1977, *84,* 405-411.

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

Montanelli, R. G., & Humphreys, L. G. Latent roots of random data correlation matrices with squared multiple correlations on the diagonal: A Monte Carlo study. *Psychometrika,* 1976, *41,* 341-348.

Revelle, W., Humphreys, M. S., Simon, L., & Gilliland, K. The interactive effect of personality, time of day and caffeine: A test of the arousal model. In press, *Journal of Experimental Psychology: General.*

Sörbom, D., & Jöreskog, K. *COFAMM.* Chicago: National Educational Resources, Inc., 1976.

Thurstone, L. L. *Multiple factor analysis.* Chicago: University of Chicago Press, 1947.