

Power as a function of communality in factor analysis

PETER H. SCHÖNEMANN
Purdue University, Lafayette, Indiana 47907

It was recently reported that the likelihood ratio test (LRT) in unrestricted factor analysis "has considerable power even when the sample size is only 10." Here it is shown that (1) the power of this LRT depends on the communalities (h^2), (2) the average h^2 in the simulation on which this conclusion was based far exceeds any found in French's review of 20 years of factor-analytic work, and (3) for more realistic h^2 , the power of the test barely exceeds the α level for $N = 30$ and remains poor for N s as large as 100.

Recently, Geweke and Singleton (1980) published the results of a Monte Carlo study designed to evaluate the performance of the likelihood ratio test (LRT) of unrestricted maximum likelihood factor analysis (UMLFA) in small samples.

The conventional factor analysis model, which goes back to Spearman (1904), can be written

$$\eta = A\xi + U\xi, \quad (1)$$

where $\eta' = (y_1, \dots, y_p)$ contains p observed variables y_i , $\xi' = (x_1, \dots, x_m)$ contains m ($1 \leq m \leq p-1$) "common factors" x_j , and $\xi' = (z_1, \dots, z_p)$ contains p "unique factors" z_i , which are assumed to jointly satisfy the covariance requirement,

$$\text{var}(\xi', \xi') = I \text{ of order } p + m. \quad (2)$$

The $p \times m$ matrix $A = (\alpha_{ij})$ is assumed to have full column rank m and is called the "common factor pattern." Equations 1 and 2 imply that it is the regression pattern of η on ξ , containing the regression weights for predicting the observed y_i from the m "latent" common factors x_j as rows. The $p \times p$ matrix U is the corresponding regression pattern for the unique factors z_j . It is assumed to be positive definite and diagonal. These definitions imply

$$\text{var}(\eta) = \Sigma = A A' + U^2, \quad (3)$$

and, further, that the partial correlation matrix,

$$\text{corr}(\eta | \xi) = U^{-1}(\Sigma - A A')U^{-1} = I_p, \quad (4)$$

is diagonal. This is a necessary condition for the factor model to hold for m common factors, and it can be tested, at least in principle, as a statistical hypothesis in applications of this model, if one appends the usual multinormality assumption (e.g., Howe, 1955; Lawley, 1940). The ratios

$$h_i^2 = \text{var}(\alpha_i' \xi) / \text{var}(y_i) \quad (5)$$

have been traditionally called "communalities." They can be interpreted as the variances of the "observed parts" $y_i = \alpha_i' \xi$ of the observed variables if the y_i have been standardized or, equivalently, as the squared multiple correlations between the observed variables y_i and the common factors x_j .

Geweke and Singleton (1980) generated two population covariance matrices $\Sigma_k = A_k A_k' + U^2$ for the two factor patterns,

$$A_1' = (4 \ 4 \ 3 \ 3 \ 2), \quad A_2' = \begin{pmatrix} 4 & 4 & 3 & 3 & 2 \\ 1 & 2 & 2 & 1 & 2 \end{pmatrix} \quad (6)$$

and the matrix of unique variances,

$$\text{diag}(U^2) = \text{diag}(2.0 \ 1.66 \ 1.33 \ .66 \ .33). \quad (7)$$

They then drew random samples of sizes $N = 10, 30, 150$, and 300 and analyzed the resulting sample covariance matrices with a maximum likelihood routine similar to that of Jöreskog (1967). On replicating the analysis for various combinations of sample sizes, the two-factor patterns, and the null hypotheses $H_0: m = 0, 1, 2$, Geweke and Singleton reached this surprising conclusion:

"First, this statistic may be more reliable in small samples than previously believed: the suggestion of Lawley and Maxwell (1971, p. 36) that Bartlett's version of the test can be trusted only if $N - p \geq 50$ is probably too pessimistic. The fewer factors being fit, the sooner the asymptotic theory becomes appropriate as sample size is increased, the threshold being approximately 10 observations for one factor and perhaps 25 for two Second, the likelihood ratio test has considerable power even when sample size is only 10" (Geweke & Singleton, 1980, p. 136).

THE COMMUNALITY PROBLEM

Geweke and Singleton (1980) deserve credit for drawing attention to the important and long-neglected power

problem in UMLFA. However, their results cannot be expected to generalize to correlation matrices that typically arise in real life, because the simulation employed a population covariance matrix with unrealistically large communalities. The correlation matrix for A_2 associated with Equation 6 and U^2 (Equation 7) is reproduced in the upper triangle of Table 1. With one exception, all correlations are in the high 80s or low 90s. Such correlations are rarely found in the social sciences, in which factor analysis is primarily used, because most measures are not reliable enough to permit such a high share of the variance. As a consequence, the communalities employed in the Geweke and Singleton (1980) study are equally unrealistic, averaging .93 (see Table 2).

A fairly representative estimate of the communality range encountered in the social sciences can be distilled from a review of some 69 factor-analytic studies conducted in the 1940s (French, 1951). The distribution of the average communalities for the 61 studies with admissible communality estimates is given in Table 3. In no case does the average communality reach the value employed in the Geweke and Singleton (1980) study. The two highest values stay well below .75, and the average is around .55. The studies reviewed by French cover a wide spectrum of ability tests, sample sizes, and extraction methods of the period, including most commercial tests in use at one time, or even now. They thus can be viewed as a representative sample of the best efforts of the factor-analytic tradition in the social sciences prior to the arrival of the presumably more sophisticated iterative algorithms developed by Browne (1969), Jöreskog (1967), Howe (Note 1), and others in the 1960s and early 1970s. Since communalities depend on the data, not on the algorithm used for their analysis, it is not likely that the average communalities have

Table 1
Correlation Matrices Implied by $\Sigma_k = A_2 A_2' + q_k U^2$

	1	2	3	4	5
1	1.00	.89	.85	.91	.80
2	.49	1.00	.91	.92	.89
3	.45	.52	1.00	.89	.92
4	.53	.57	.53	1.00	.85
5	.45	.59	.58	.58	1.00

Note—Upper triangle: $q_k = 1.0$ (Geweke & Singleton, 1980, study). Lower triangle: $q_k = 10$ (approximate average of French, 1951, survey).

Table 2
Communalities Implied by Various Multipliers q_k

q_k	1	2	3	4	5	h^2	$c(k)_2$
1.0	.90	.92	.91	.94	.96	.93	2.92*
1.5	.85	.89	.87	.91	.94	.89	1.95
2.0	.81	.86	.83	.88	.92	.86	1.47
3.0	.74	.80	.77	.84	.89	.81	.98
5.0	.63	.71	.66	.75	.83	.72	.59
7.0	.55	.63	.58	.68	.78	.65	.42
10.0	.46	.55	.49	.60	.71	.56	.29†

*Geweke and Singleton (1980).

†French (1951) survey.

Table 3
Average Communalities in 61 Studies Reviewed by French (1951)

Upper Boundary	.35	.40	.45	.50	.55	.60	.65	.70	.75
Class Frequency	02	04	03	12	19	10	07	02	02

changed appreciably since the French study was compiled. French's review dealt with ability and achievement measures, which are among the better, more reliable measures in the social sciences. For personality measures, which have also been extensively factor analyzed, the correlations and communalities are more likely to average in the 30s and low 40s.

POWER ESTIMATES AS A FUNCTION OF COMMUNALITY

To obtain an idea of what the power of the LRT might be for the factor pattern A_2 with communalities more in keeping with the typically encountered range, random samples were drawn from multinormal populations with covariance matrices,

$$\Sigma_k = A_2 A_2' + q_k U^2, \quad (8)$$

for various values of the multiplier q_k , as given in Table 4, which thus controls the effective average communality in the population (see Table 2). Two hundred such samples were reanalyzed for $q_k = 1$ (i.e., in exact accord with the population matrix selected by Geweke & Singleton, 1980) under $H_0: m = 2$, to estimate the null distribution of the test statistic for A_2 , which has rank 2. The rest of the simulation dealt with the (incorrect) hypothesis

$$H_0: m = 1, \quad (9)$$

to assess the power of the test as a function of the average communality when A_2 is kept fixed. In all cases but one, the sample size was fixed at $N = 30$. The number of replications run for each q_k was 200. The LRT statistic employs the correction recommended by Box (1949). The UMLFA algorithm was written by Browne (1969).

The results of this replication of the Geweke and Singleton (1980) study for varied communality parameters are tabulated in Table 4 as cumulative proportions of the probabilities associated with the chi-square statistics obtained on reanalysis of the sampled correlation matrices. The empirical distributions are presented explicitly for 100 replications per row, to enable the reader to estimate the power directly and, at the same time, to convey an impression of the stability of such empirical estimates for 100 replications, the number employed in the Geweke and Singleton study.

The null distribution (first two rows of Table 4) corroborates the results obtained by Geweke and Singleton (1980) that "the null hypothesis . . . will be rejected too often" (p. 136) when it is true. They observed 13% rejections of H_0 at the .05 level for $N = 10$. In the

Table 4
Performance of LRT as a Function of Average Communality

			α Level									
q_k	h^2	St	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
m = 2, $m_0 = 2$, N = 30 (Null Distribution)												
1.0	.93	A	.16	.23	.37	.51	.67	.78	.84	.89	.94	1.0
		B	.11	.24	.35	.49	.67	.80	.83	.89	.94	1.0
m = 2, $m_0 = 1$, N = 30 (Power Estimates)												
1.0	.93	A	.79	.90	.94	.98	.99	1.00	1.00	1.00	1.00	1.0
		B	.81	.90	.95	.97	.99	.99	.99	.99	.99	1.0
1.5	.89	A	.55	.72	.79	.85	.88	.91	.95	.98	1.00	1.0
		B	.48	.65	.75	.82	.90	.94	.99	.99	1.00	1.0
2.0	.86	A	.45	.64	.72	.80	.87	.91	.91	.97	.98	1.0
		B	.42	.53	.64	.72	.80	.82	.88	.92	.96	1.0
3.0	.81	A	.22	.32	.44	.52	.62	.69	.77	.83	.90	1.0
		B	.34	.52	.62	.71	.82	.89	.90	.95	.97	1.0
5.0	.72	A	.19	.29	.42	.51	.64	.71	.72	.84	.95	1.0
		B	.19	.27	.37	.47	.59	.70	.81	.88	.95	1.0
7.0	.65	A	.14	.26	.37	.40	.52	.64	.72	.80	.93	1.0
		B	.15	.26	.38	.52	.68	.72	.82	.89	.92	1.0
10.0	.56	A	.09	.22	.36	.41	.51	.61	.72	.82	.92	1.0
		B	.17	.24	.35	.41	.49	.61	.72	.80	.95	1.0
m = 2, $m_0 = 1$, N = 100												
10.0	.56	A	.17	.32	.47	.50	.55	.62	.71	.77	.88	1.0
		B	.14	.26	.38	.52	.62	.68	.75	.87	.99	1.0

Note—Entries under α levels indicate proportions of rejections of $H_0: m = 1$. A/B = two separate studies, based on 100 replications each. N = sample size. Example: The last two rows for N = 30 give rejection proportions for population covariance matrix with average communality $h^2 = .56$ (modal value of French, 1951, review). A .1 level test rejects the (false) $H_0: m = 1$ in nine responses, or 17% of the cases, so that the estimated power of a .1 level test against the correct alternative $H_1: m = 2$ is .13.

present study, the observed frequencies of rejection were 16% and 11%, respectively, at the .1 level for N = 30.

The third and fourth rows of Table 4 give the power estimates for the same population covariance matrix for N = 30. Geweke and Singleton (1980) found that "for 30 observations the . . . proportions [of rejecting H_0 at the .05 and .01 levels] were 63 and 42 percent, and for 150 and 300 observations the one factor model was always rejected at both significance levels" (p. 136). Superficially, the present reanalysis also seems to corroborate this earlier finding, because the power at the .1 level was estimated as .80 and .79 for each of the two runs with 100 replications.

However, as the remainder of Table 4 shows, this impressive result is tied to the unrealistic choice of the population covariance matrix used by Geweke and Singleton (1980). Their results, therefore, do not warrant the following unqualified recommendation: "Based on the experimental results reported in Table 1, it seems advisable to adopt smaller significance levels in smaller samples" (Geweke & Singleton, 1980, p. 136). As can be seen from the columns of Table 4, the power fades rapidly as the average communality approaches a more realistic range. When this value attains .72, which is still at the upper tail of the distribution for the French (1951) summary, the power estimates for .1 level tests

are in the upper 10s or low 20s. For the modal value of the French distribution ($h^2 = .55$), the power barely exceeds the chosen level of significance. As the last two rows in Table 4 show, this poor performance of the test of fit improves only slightly when the sample size is raised from 30 to 100. At this point, a .2 level test has one chance out of three to detect that $H_0: m = 1$ is false.

Although these empirical results are conditional on the particular factor pattern A_2 selected by Geweke and Singleton (1980), they raise sufficient doubts about the value of the LRT in factor analysis in general to call for further, more representative investigations of the merits of maximum likelihood algorithms for factor analysis, which, especially in view of the considerable cost of these programs, are long overdue.

DISCUSSION

Although the power estimates in Table 4 are conditional on the particular factor pattern A_2 selected by Geweke and Singleton (1980), the monotonic relationship between power and average communality will hold in general. To see this, let

$$E = U^{-1} (C - A A') U^{-1} \quad (10)$$

be the sample analog of the partial correlation matrix (Equation 4), where C is the observed covariance matrix, and U^2 and A are the maximum likelihood estimates (LMEs) of U^2 and A. Then the LRT statistic can be written

$$T(Y) = -\text{const} \ln |E|. \quad (11)$$

This criterion can be viewed as a measure of the departure of the latent roots of E around their average, unity. If one denotes the eigen decomposition of $U^{-1} C U^{-1}$

$$U^{-1} C U^{-1} = L(D + I)L', \quad (12)$$

where

$$D + I = \text{diag}(c_1 + 1, \dots, c_p + 1) \quad (13)$$

contains the p latent roots, ordered by size, L contains the associated latent vectors, and L_1 is the pxm matrix of latent vectors, which correspond to the m largest latent roots $c_1 + 1, \dots, c_m + 1$, then the MLE of A is

$$A = U L_1 D_1^{-1/2}, \quad (14)$$

and it will satisfy the identifiability condition

$$A' U^{-2} A = \text{diag}(c_1, \dots, c_m). \quad (15)$$

Since the two matrices on the right of Equation 10 share the same eigen vectors L_1 , Equation 11 can be written in terms of the c_i

$$T(Y) = -\text{const} [\ln(1 + c_{m+1}) + \dots + \ln(1 + c_p)]. \quad (16)$$

If the (m + 1)st c_j is close to zero, it will not contribute much to this measure of fit, and the test will tend to accept H_0 for $m_0 < m$. But if c_j^2 is the latent roots of $A' U^2 A$ associated with $\Sigma_1 = A A' + U^2$, then

$$c^{(k)}_j = c_j / q_k \quad (17)$$

is the corresponding roots of $A' U^{(k)2} A$ associated with $\Sigma_k =$

$AA' + U^{(k)2} = AA' + q_k U^2$. As an illustration, the m 'th (second) root for the Σ_k of the present study is given in the last column of Table 2. For a suitably large q_k , the m 'th root can be made as small as one pleases. Although A will retain full column rank m for all finite choices of q_k , this fact will be more and more difficult to detect in the sample with the measure of fit (Equation 16). As q_k increases and, consequently, h^2 and $c^{(k)}_m$ decrease, a test of $H_0: m_0 < m$ will lose power against the true alternative $H_1: m_0 = m$.

The m 'th root c_m is also a critical indicator of the definition of the common factors. If this root is small, then the m 'th common factor will be poorly defined. This peculiar property of the conventional factor model (Equations 1 and 2), that it allows for the existence of other sets of common factors, x_j^* , when at least one such set, x_j , exists, was first noticed by Wilson (1928/ in press). For an account of the uneven history of this indeterminacy problem, see Steiger and Schönemann (1978); for some of the implications, see Schönemann and Steiger (1978) and Steiger (1979).

The minimum correlation between two equivalent common factors, x_j and x_j^* , can be expressed as

$$\rho_j = (c_j - 1)/(c_j + 1), j = 1, \dots, m \quad (18)$$

(Schönemann & Wang, 1972). If $c_j < 1$, then two equivalent common factors x_j and x_j^* exist that are negatively correlated. Schönemann and Wang found that such negative correlations arise in practice quite often once m is raised to achieve a statistical fit.

As long as the communalities of the actual test batteries remain in the range indicated by the French (1951) review, the m 'th root c_m is likely to be small, and, as a result, users of the factor model will be confronted with two problems at once: poorly determined factors and low power of the tests of fit. Both problems can be traced algebraically to the magnitude of the smallest root c_m in the population.

REFERENCE NOTE

1. Howe, W. G. *Some contributions to factor analysis* (Report

No. ORNL-1919). Oak Ridge, Tenn: Oak Ridge National Laboratory, 1955.

REFERENCES

- Box, G. E. P. A general distribution theory for a class of likelihood criteria. *Biometrika*, 1949, **36**, 317-346.
- Browne, M. W. Fitting the factor analysis model. *Psychometrika*, 1969, **34**, 375-394.
- French, J. W. The description of aptitude and achievement tests in terms of rotated factors. *Psychometric Monographs*, 1951, **5**, 000-000.
- Geweke, J. F., & Singleton, K. J. Interpreting the likelihood ratio statistic in factor models when sample size is small. *Journal of the American Statistical Association*, 1980, **75**, 133-137.
- Jöreskog, K. G. Some contributions to factor analysis. *Psychometrika*, 1967, **32**, 443-482.
- Lawley, D. N. The estimation of factor loadings by the method of maximum likelihood. *Proceedings of the Royal Society of Edinburgh* (Section A), 1940, **60**, 64-82.
- Lawley, D. N., & Maxwell, A. E. *Factor analysis as a statistical method*. New York: American Elsevier, 1971.
- Schönemann, P. H., & Steiger, J. H. On the validity of indeterminate factor scores. *Bulletin of the Psychonomic Society*, 1978, **12**, 287-290.
- Schönemann, P. H., & Wang, M. M. Some new results on factor indeterminacy. *Psychometrika*, 1972, **37**, 61-91.
- Spearman, C. General intelligence, objectively determined and measured. *American Journal of Psychology*, 1904, **15**, 201-293.
- Steiger, J. H. The relationship between external variables and common factors. *Psychometrika*, 1979, **44**, 193-197.
- Steiger, J. H., & Schönemann, P. H. A history of factor indeterminacy. In S. Shye (Ed.), *Theory construction and data analysis in the social sciences*. San Francisco: Jossey Bass, 1978.
- Wilson, E. B. Review of *The abilities of man, their nature and measurement*, by C. Spearman. In I. Borg (Ed.), *Multidimensional data representation: When and why*. Ann Arbor: Mathesis Press, in press. (Reprinted from *Science*, 1928, **67**, 244-248.)

(Received for publication November 12, 1980.)