PSYCHOMETRIKA—VOL. 76, NO. 3, 461–470
JULY 2011
DOI: 10.1007/S11336-011-9217-5

FACTOR ANALYSIS VIA COMPONENTS ANALYSIS

PETER M. BENTLER AND JAN DE LEEUW

UNIVERSITY OF CALIFORNIA, LOS ANGELES

When the factor analysis model holds, component loadings are linear combinations of factor loadings, and vice versa. This interrelation permits us to define new optimization criteria and estimation methods for exploratory factor analysis. Although this article is primarily conceptual in nature, an illustrative example and a small simulation show the methodology to be promising.

Key words: factor loadings, factor scores, component loadings, component scores.

Although limiting conditions have been developed under which components, or principal components as an important special case, and latent factors of factor analysis coincide (e.g., Guttman, 1956; Bentler & Kano 1990), in most treatments components analysis and factor analysis are considered to be alternative but basically unrelated methods for determining sources of variance in variables (e.g., Mulaik, 2010). The goal of this paper is to estimate the parameters of the common factor model via components. To do this, we require a stronger linking between the models than has been previously described. To begin, we provide some matrix background on the existence of a factor analysis model.

Suppose C is a symmetric positive semidefinite matrix of order p and rank r with $C = U\Delta^2U'$, where U is a $p \times r$ orthonormal matrix of eigenvectors with $U'U = I$, and Δ^2 is an $r \times r$ diagonal matrix of eigenvalues. Then we have

Result 1. A $p \times q$ matrix X satisfies $C = XX'$ if and only if $X = U\Delta V'$, where V is $q \times r$ and satisfies $V'V = I$. Thus q cannot be smaller than r, and the rank of X is equal to r.

Postmultiplying $X = U\Delta V'$ by its transpose gives C, which is of rank r, establishing that $q \geq r$ and the rank of X. Assuming to the contrary that $V'V \neq I$ violates the assumption that $C = U\Delta^2U'$, thus leading to a contradiction. Note that $X = U\Delta V'$ gives its singular value decomposition.

Now let us take $C = \Sigma$, a population covariance matrix with eigenvector decomposition $\Sigma = U\Delta^2U'$. The orthogonal factor analysis model states that the covariance matrix has decomposition $\Sigma = \Gamma\Gamma'$ where $\Gamma = [\Lambda|\Psi]$ is a $p \times q$ $(= k + p)$ partitioned matrix of factor loadings that contains a $p \times k$ common factor loading matrix Λ and the $p \times p$ diagonal unique loading matrix Ψ. We apply Result 1 to the factor model by letting $X = \Gamma$, obtaining

Result 2. The orthogonal factor analysis model $\Sigma = \Gamma\Gamma'$ is true if and only if there exists $\Gamma = [\Lambda|\Psi]$ with Λ, diagonal Ψ, and orthonormal matrix V such that $\Gamma = U\Delta V'$.

It follows that $\Gamma V_\perp = 0$, where $V_\perp$ is a $q \times (q - p)$ orthogonal complement such that $V_\perp' V = 0$. In the following we assume that the factor analysis model $\Sigma = \Lambda\Lambda' + \Psi^2$ holds in the population.

Requests for reprints should be sent to Peter M. Bentler, Departments of Psychology and Statistics, UCLA, Box 951563, Los Angeles, CA 90095-1563, USA. E-mail: bentler@ucla.edu

1. Interrelations Between Models

Under a component model, a random p-variate vector of observed variables x may be expressed as a population model based on a linear combination of components ζ with coefficients L

$$x = L\zeta. \tag{1}$$

L is typically called a component loading matrix and is sometimes misleadingly called a factor loading matrix. We take L to be square and full rank, so that this is a complete components representation.[1] It is not necessarily unique, since (1) allows a rotation. One way to make it unique is to specify that $L'L$ is diagonal with elements ordered from large to small, i.e., a principal components representation. In practice, when the decomposition (1) is applied with sample data, it is used as an approximation, and the number of columns of L is taken to be substantially below p.

Under a common factor analysis model, the same p-variate vector of observed variables x is given a different population decomposition. In particular, using the factor notation above,

$$x = \Gamma\xi = [\Lambda|\Psi]\xi \tag{2}$$

where the q random variables ξ are factors consisting of k common factors and p unique factors. Now, accepting Result 2, we assume that both models (1) and (2) are true in the population. In this case, we may write $L\zeta = \Gamma\xi$, and, since the complete component loading matrix L is invertible, we have

$$\zeta = L^{-1}\Gamma\xi. \tag{3}$$

That is, we have the obvious result.

Result 3. Components ζ are linear combinations of factors ξ.

In particular, components are combinations of common and unique factors. Since unique factors contain specificity plus random error, components also contain these sources of variance.

Next we consider the covariance structures of these two models. Under the usual assumptions that $E(x) = 0$, $E(\zeta) = 0$, $E(\xi) = 0$, the components ζ are mutually uncorrelated, and, in exploratory factor analysis, the factors ξ are mutually uncorrelated, when both models are true we have

$$\Sigma = LL' = \Gamma\Gamma'. \tag{4}$$

Now we perform the singular value decomposition given in Result 2, namely, $\Gamma = U\Delta V'$, where U is $p \times p$ with $U'U = I$, Δ is a $p \times p$ diagonal matrix of singular values, and V' is $p \times q$ with $V'V = I$. Furthermore, since $\Gamma VV' = \Gamma$, we may write $LL' = \Gamma VV'\Gamma'$. Since we have specified no special structure for L, and since ΓV is of dimension $p \times p$, this allows us to take

$$L = \Gamma V. \tag{5}$$

Furthermore, it follows trivially from the singular value decomposition of Γ and (5) that

$$\Gamma = LV'. \tag{6}$$

Thus we have

[1] If L is not full rank, the population covariance matrix is singular and the factor analysis model $\Sigma = \Lambda\Lambda' + \Psi^2$ cannot hold, violating our assumption.

Result 4. The component loading matrix L is a rank-reducing linear combination of elements of the factor loading matrix Γ.

Suppose D_p is a $p \times p$ nonsingular diagonal matrix, and that the observed variables are scaled by this matrix. The effect on the component loadings is to yield $L_D = D_p L$, and on the factor loadings is to yield $\Gamma_D = D_p \Gamma$. Clearly, these rescaled loading matrices maintain the relations given by (5) and (6), and hence we may, without loss of generality, perform all analyses on the correlation matrix. Thus we have

Result 5. The relations between component and factor loadings given by (5) and (6) are invariant to rescaling of the observed variables.

Earlier we stated that the choice of component representation in the above relations is arbitrary. It may be worthwhile to be explicit about why this is so. Consider transforming the component loading matrix by an orthonormal matrix. If T is a matrix such that $T'T = TT' = I$, and the left- and right-hand sides of (5) are postmultiplied by T, the relation (5) is maintained for the new loading matrix $L_T = LT$ and the new $V_T = VT$, where V_T possesses the same orthogonality properties as the original V. Hence

Result 6. The choice of components, such as principal components, is arbitrary. Any convenient component representation maintains the key relations between components and factors.

It may be useful to give the explicit interrelations between components and factors at the level of population covariances. We postmultiply (3) by ξ' and take expectations of both sides, yielding

$$\Phi_{\zeta\xi} = E(\zeta\xi') = L^{-1}\Gamma. \tag{7}$$

If the starting point of our analysis had allowed the factors to be correlated with covariance matrix Φ, this covariance matrix would appear on the right side of (7). In any case, the interpretation of elements of $\Phi_{\zeta\xi}$ will hinge critically on the chosen scaling of variables, the identification conditions utilized, and possible rotations imposed on the components as well as the factors. Thus we have

Result 7. The covariance matrix relating components and factors, or their correlation matrix when standardized, is given by $L^{-1}\Gamma$.

The above equations are population relations that hold if the hypothesized factor model is true. Otherwise, they will only be approximations, and the quality of the approximation will depend on the correctness of (2), as well as a correct choice of the number of factors. For example, if (2) holds with k common factors, a factor model with $k-1$ factors would not be consistent with (5) and (6) in either the population or in a sample. Equations (5) and (6) would be approximations rather than equalities. In addition, when applied to real data, where a sample covariance matrix S will replace its corresponding population Σ to yield a sample component loading matrix $\hat{L}$, (5) and (6) will no longer hold exactly. This implies the need for a methodology to optimize the approximations.

2. Practical Approaches to Factor Analysis via Components

The interrelations developed above can be used in several ways to obtain new factor analytic estimation methods. The most obvious approach is to define optimization functions based on the population relations (5) and (6), and then to apply them to parameter estimation with sample data. In the population, minimized values of these functions provide sample-size independent definitions of population lack of fit when the chosen factor model is not true; whereas when implemented with sample components, they define discrepancy functions to be minimized in an estimation methodology. Because of the scale invariance noted in Result 5, we can work with the correlation matrix without loss of generality, i.e., we may consider the above relations based on standardized observed variables. That is, we consider $P = LL'$, where P is the population correlation matrix.

Some interesting population functions can be defined on the discrepancies between the left and right sides of (5) and (6). In this paper, we consider only relatively simple nonnegative functions that take on the value of zero under the null hypothesis (2). First we consider a kind of generalized least squares function based on (5) that, in contrast to previous approaches to factor analysis that fit the factor model to the correlation matrix, fits the factor model to the component loading matrix

$$\delta_1 = \frac{1}{2}\,\mathrm{tr}(L - \Gamma V)' W (L - \Gamma V) \quad \text{with } \Gamma V_\perp = 0. \tag{8}$$

W is a weight matrix that can be chosen in various ways. When $W = P^{-1}$, (8) is a scale-invariant way to measure how close $\Gamma V L^{-1}$ is to an identity matrix based on $\delta_1 = .5\,\mathrm{tr}(I - \Gamma V L^{-1})'(I - \Gamma V L^{-1})$ and $\Gamma V_\perp = 0$. Although here we use a trace operator to quantify this, it also is possible to use $|\Gamma V L^{-1}|$ alone or in combination with (8). When $W = I$ this is a least squares (LS) discrepancy function.[2] When a sample $\hat{L}$ is fitted, the formulation allows such options as $W = R^{-1}$, based on the sample correlation matrix R, and $W = \hat{P}^{-1}$ based on the estimated model $\hat{P} = \hat{\Gamma}\hat{\Gamma}'$; these may be called generalized least squares (GLS) and reweighted least squares (RLS) methods, respectively.

In parallel to the above, we also may consider a function based on (6) that fits the factor loading matrix to a transformed component loading matrix, specifically

$$\delta_2 = \frac{1}{2}\,\mathrm{tr}\big(\Gamma - LV'\big)' W \big(\Gamma - LV'\big). \tag{9}$$

In this case, when $W = P^{-1} = (LL')^{-1} = (\Gamma\Gamma')^{-1}$, the function simplifies to the discrepancy $\delta_2 = p - \mathrm{tr}(\Gamma V L^{-1})$, and as with (8), alternative specifications of W lead to LS-, GLS-, and RLS-type methods. Of course, additional methods are possible as noted in the Discussion.

Although there are many options for estimation, a careful study of alternatives is beyond the scope of this paper. We may call this class of procedures exploratory factor analysis via components (EFAC). We can show, using Slutsky (1925) in the same way as Theorem 1 of Shapiro (1984), that EFAC estimators of factor loadings and uniquenesses obtained by minimizing (9) [or (8)] are consistent estimators of their population counterparts. This is true for both the random factor score as well as the fixed factor score models, provided the sample covariance matrix converges in probability to the population covariance matrix. Since the minimizer of the loss function is a continuous function of the covariance matrix, and since for the population covariance matrix the loss function is minimized at the population loadings and uniquenesses, we obtain the required consistency result. Note that this result holds for any identified component representation, i.e., it is true when L and its sample counterpart $\hat{L}$ are principal components, a Cholesky decomposition, or the symmetric square root. We simply illustrate the proposed approach with one method and one data set, and a small simulation, taking $L = \hat{L}$ from $R = \hat{L}\hat{L}'$ and treating it as fixed.

[2] Least squares (LS) is sometimes called unweighted least squares (ULS) or ordinary least squares (OLS).

3. Computational Approach

We illustrate our results by minimizing (9) with $W = I$, i.e., using the least squares special case. What is different in our approach is that V is also an unknown matrix that needs to be estimated along with Γ, although it will typically not be of special interest. Here we consider a simple alternating least squares approach to obtain the unknown optimum parameter estimates $\hat{\Gamma}$ and $\hat{V}$. We use two steps:

1. Given a current Γ, say Γ_i, find a V, say V_i;
2. Given a current V_i, find a new Γ, say Γ_{i+1},

repeated in sequence until convergence. Abstractly, step 1 is implemented by obtaining the gradient $\partial\delta_i/\partial V$ and computing up a type of gradient projection step (Jennrich, 2002) involving a least squares orthonormalization similar to that of orthogonal Procrustes rotation (Schönemann, 1966). This makes use of a Lagrangian constraint to ensure that V is an orthonormal matrix at each iteration, and ensures that $\Gamma V_\perp = 0$. Step 2 is implemented by a similar gradient projection step that ensures that the estimated $\hat{\Gamma}$ has the required form (2). Also, to disallow rotation and column sign changes in Λ, we define the population Λ as having an upper-right triangle of fixed zeros with the largest element in each column as positive. We impose the same conditions on Λ_i at each iteration, and hence on the final $\hat{\Lambda}$.

Specifically, in step 1, using the current estimates we compute the steepest-descent update matrix on the left of the equality

$$V - \alpha\big(VL'L - \Gamma'L\big) = PDQ', \tag{10}$$

and obtain its singular value decomposition as given on the right of (10). Here, α is a step-size coefficient chosen to guarantee a decrease in function (9); Jennrich proved that such a step size exists. The usual properties ensure that $P'P = I$, $Q'Q = I$, and D is the diagonal matrix of singular values. Then the new estimate

$$V_i = PQ' \tag{11}$$

is obtained. This is the least squares orthonormal approximation to the matrix in (10). This new estimate V_i is used in step 2, which proceeds by computing $\partial\delta_{2\text{LS}}/\partial\Gamma_f = 0 \Rightarrow \Gamma_{i+1} = [LV_i']_f$ where Γ_f is the vector of free parameters in Γ and $[.]_f$ extracts the elements corresponding to free parameters in Γ and ignores the rest. Then we return to step 1 and continue cycling until a minimum of (9) is obtained, yielding the final $\hat{\Gamma}$ and $\hat{V}$.

4. Example

To illustrate this methodology, we analyzed the well-known Holzinger–Harman 24 psychological tests (Harman, 1976) and compared the EFAC results to maximum likelihood (ML). Convergence to the minimum of (9) was straightforward. Both results were rotated by varimax; Table 1 gives the two solutions. Even though the use of separate rotations for the two solutions does not maximize the similarity of solutions, as a target rotation might, the EFAC solution on the left side of Table 1 is remarkably similar to the ML solution given on the right side of Table 1: the root mean square difference across all elements of the two matrices is 0.012. A similar comparison of the EFAC solution to the standard iterative principal axes solution (not shown) obtained a root mean square difference of 0.009.

TABLE 1.
EFAC (*left*) and ML (*right*) varimax rotated solutions for 24 psychological variables.

1	0.153	0.704	0.169	0.130	0.160	0.689	0.187	0.161
2	0.119	0.435	0.099	0.081	0.117	0.436	0.083	0.096
3	0.141	0.534	−0.010	0.159	0.137	0.571	−0.020	0.109
4	0.227	0.553	0.088	0.054	0.233	0.528	0.099	0.079
5	0.741	0.185	0.216	0.148	0.739	0.185	0.213	0.150
6	0.764	0.209	0.067	0.232	0.767	0.205	0.066	0.233
7	0.809	0.201	0.152	0.070	0.806	0.197	0.153	0.075
8	0.569	0.342	0.233	0.139	0.569	0.339	0.242	0.131
9	0.810	0.209	0.043	0.215	0.806	0.201	0.040	0.227
10	0.168	−0.098	0.824	0.157	0.168	−0.118	0.831	0.167
11	0.176	0.111	0.530	0.383	0.179	0.119	0.511	0.378
12	0.019	0.209	0.719	0.087	0.019	0.210	0.716	0.089
13	0.181	0.428	0.535	0.084	0.187	0.437	0.525	0.083
14	0.202	0.045	0.081	0.574	0.197	0.050	0.081	0.554
15	0.120	0.125	0.077	0.514	0.121	0.116	0.075	0.522
16	0.068	0.424	0.053	0.517	0.069	0.408	0.062	0.525
17	0.138	0.069	0.221	0.583	0.142	0.062	0.219	0.573
18	0.022	0.308	0.339	0.447	0.026	0.294	0.336	0.455
19	0.145	0.244	0.167	0.364	0.148	0.240	0.161	0.365
20	0.378	0.421	0.107	0.288	0.378	0.402	0.118	0.300
21	0.175	0.407	0.433	0.200	0.175	0.381	0.438	0.223
22	0.365	0.414	0.125	0.283	0.366	0.399	0.122	0.301
23	0.370	0.517	0.234	0.224	0.369	0.501	0.244	0.238
24	0.367	0.179	0.492	0.294	0.370	0.158	0.496	0.303
Col SSQ	3.647	2.975	2.657	2.257	3.647	2.875	2.654	2.292

TABLE 2.
Population factor loading matrix for simulation.

0.7	0	0
0.8	0.5	0
0.6	0.7	0.3
0.5	0.2	0.2
0.6	0.3	0.3
0.2	0.6	0.1
0.3	0.8	0.4
0.1	0.6	0.3
0.2	0.4	0.1
0.1	0.5	0.1
0.2	0.4	0.8
0.3	0.1	0.9
0.2	0.3	0.7
0.4	0.4	0.5
0.1	0	0.5

5. Simulation

A small simulation was undertaken to provide further provisional information on the feasibility of the EFAC methodology. The 15 by 3 population factor loading matrix Λ shown in Table 2, with its upper-right triangle of zeros, was used to generate a population correlation matrix P with $\Psi^2 = P - \Lambda\Lambda'$. Five hundred random normal samples of size 1000 were taken from this population. In each sample, the sample correlation matrix was computed and factor analyzed by EFAC. The procedure was repeated with ML estimation. Further, the methodology was repeated using sample size 300. Under all conditions, in each replication the estimated factor loading matrix was taken as lower triangular, so that each $\hat{\Lambda}$ directly estimated the 42 free parameters corresponding to the lower triangle of Λ without the need for further rotation.

The simulation results are summarized in Table 3 for $N = 1000$ and in Table 4 for $N = 300$. The first column in these tables provides a label for each factor loading of Table 2, with *Vi*, *Fj* referring to the loading of the ith variable on the jth factor. The accuracy of each factor loading estimate on average is given in columns 2 and 3, which report the differences between the true parameter value (Table 2) and its mean estimate across the 500 replications. Column 2 shows the algebraic difference between each mean EFAC estimate and its corresponding true value; column 3 shows this difference for the ML estimates; and column 4 shows the differences between mean ML and EFAC estimates.

At $N = 1000$ (Table 3), the average EFAC estimate equals its population value to the third decimal in precision. The root mean square (RMS) difference between the 42 mean EFAC estimates and their population counterparts is 0.0027. Similar results were obtained for ML, though its RMS difference is slightly smaller at 0.0021. Further, the EFAC and ML estimates of the individual parameters are remarkably similar, differing only in the third decimal. They are equal up to an RMS difference of 0.0023.

The results at $N = 300$ (Table 4) similarly show that EFAC produces factor loadings that are consistent for their population counterparts, although precision is somewhat reduced due to the smaller sample size. The several RMS values are roughly twice as large as their $N = 1000$ counterparts. The RMS difference for EFAC is 0.0067, 0.0046 for ML, and 0.0043 between EFAC and ML. While EFAC performed very well at recovering the population parameters, ML was again slightly more accurate. Yet differences between EFAC and ML estimates only showed up in the third decimal place, on average.

6. Discussion

As far as can be determined, the interrelations between components and latent factor scores as well as between components and factor loading matrices that were developed in this article have not previously been recognized. The results provide the basis for a class of new estimation methods, called EFAC, for the century-old factor analysis model. In this approach, the components loading matrix rather than the correlation matrix is used to define a discrepancy function. The illustrative results on the classical 24 Holzinger–Harman psychological variables verify that an EFAC solution can be surprisingly similar to the classical maximum likelihood and least squares solutions, even after independent rotations. More importantly, the simulation verified our theoretical consistency result, showing that EFAC estimates are consistent for their population parameters. Further, EFAC was shown to yield results that are trivially different from ML estimates. This suggests that further research into the properties of EFAC may be of interest in the future.

Being primarily conceptual in nature, this paper has not addressed the wide variety of estimation methods that can be developed based on this conceptualization. The functions (8) and (9) are just illustrative, and many variants are possible. For example, δ_2 is based on a $p \times p$ weight matrix W, but the variant $\delta_2^* = \frac{1}{2}\mathrm{tr}(\Gamma - LV')W(\Gamma - LV')'$ would be based on a $q \times q$ weight matrix. Or a discrepancy function could be based on the differences between the singular values σ_i of L and singular values Δ_i of Γ, such as $\sum(\sigma_i - \Delta_i)^2$. This is an adaptation of a criterion introduced by de Leeuw (2004), the sum of squared differences between singular values of a

TABLE 3.
Simulation results for $N = 1000$.

	EFAC-TRUE	ML-TRUE	ML-EFAC
V1, F1	0.0042	0.0013	−0.0029
V2, F1	−0.0050	−0.0028	0.0022
V2, F2	0.0055	0.0026	−0.0029
V3, F1	−0.0048	−0.0015	0.0033
V3, F2	0.0037	0.0007	−0.0030
V3, F3	−0.0003	−0.0029	−0.0026
V4, F1	−0.0008	−0.0007	0.0001
V4, F2	0.0026	0.0021	−0.0005
V4, F3	−0.0008	−0.0010	−0.0002
V5, F1	−0.0022	−0.0018	0.0004
V5, F2	0.0026	0.0027	0.0001
V5, F3	−0.0011	−0.0021	−0.0010
V6, F1	−0.0020	0.0002	0.0022
V6, F2	−0.0004	−0.0010	−0.0006
V6, F3	0.0011	−0.0018	−0.0029
V7, F1	0.0009	0.0005	−0.0004
V7, F2	−0.0069	−0.0018	0.0051
V7, F3	0.0009	−0.0031	−0.0040
V8, F1	−0.0056	−0.0038	0.0018
V8, F2	−0.0017	−0.0009	0.0008
V8, F3	0.0013	−0.0018	−0.0031
V9, F1	−0.0033	−0.0017	0.0016
V9, F2	−0.0013	−0.0020	−0.0007
V9, F3	−0.0009	−0.0025	−0.0016
V10, F1	0.0002	0.0026	0.0024
V10, F2	−0.0003	−0.0013	−0.0010
V10, F3	0.0016	−0.0007	−0.0023
V11, F1	0.0010	0.0006	−0.0004
V11, F2	−0.0028	0.0017	0.0045
V11, F3	−0.0007	−0.0041	−0.0034
V12, F1	0.0006	0.0000	−0.0006
V12, F2	0.0007	0.0029	0.0022
V12, F3	−0.0057	−0.0031	0.0026
V13, F1	0.0012	0.0009	−0.0003
V13, F2	−0.0032	0.0004	0.0036
V13, F3	−0.0015	−0.0037	−0.0022
V14, F1	−0.0004	0.0002	0.0006
V14, F2	−0.0014	0.0002	0.0016
V14, F3	−0.0010	−0.0027	−0.0017
V15, F1	0.0015	−0.0001	−0.0016
V15, F2	0.0005	0.0040	0.0035
V15, F3	0.0003	−0.0007	−0.0010

raw score data matrix and those of Γ. Especially interesting would be the development of discrepancy functions that provide the type of statistics of typical interest in covariance structure analysis, such as goodness-of-fit χ^2 tests and standard error estimates for the parameters. These might include such obvious variants as asymptotically distribution free methods, and elliptical, normal theory, and robust methods based on the distribution of $\hat{L}$ and the data.

TABLE 4.
Simulation results for $N = 300$.

	EFAC-TRUE	ML-TRUE	ML-EFAC
V1, F1	0.0150	0.0060	−0.0090
V2, F1	−0.0118	−0.0046	0.0072
V2, F2	0.0109	0.0026	−0.0083
V3, F1	−0.0144	−0.0063	0.0081
V3, F2	0.0063	−0.0014	−0.0077
V3, F3	0.0021	−0.0014	−0.0035
V4, F1	−0.0048	−0.0031	0.0017
V4, F2	0.0045	0.0014	−0.0031
V4, F3	−0.0027	−0.0024	0.0003
V5, F1	−0.0045	−0.0024	0.0021
V5, F2	0.0016	−0.0009	−0.0025
V5, F3	−0.0020	−0.0030	−0.0010
V6, F1	−0.0056	0.0001	0.0057
V6, F2	−0.0018	−0.0049	−0.0031
V6, F3	−0.0027	−0.0060	−0.0033
V7, F1	−0.0072	−0.0053	0.0019
V7, F2	−0.0121	−0.0051	0.0070
V7, F3	0.0018	−0.0036	−0.0054
V8, F1	−0.0067	−0.0014	0.0053
V8, F2	−0.0058	−0.0049	0.0009
V8, F3	−0.0022	−0.0069	−0.0047
V9, F1	−0.0069	−0.0034	0.0035
V9, F2	−0.0004	−0.0023	−0.0019
V9, F3	−0.0007	−0.0017	−0.0010
V10, F1	−0.0076	−0.0036	0.0040
V10, F2	−0.0074	−0.0084	−0.0010
V10, F3	0.0013	−0.0017	−0.0030
V11, F1	−0.0055	−0.0042	0.0013
V11, F2	−0.0096	−0.0039	0.0057
V11, F3	−0.0027	−0.0083	−0.0056
V12, F1	−0.0070	−0.0062	0.0008
V12, F2	−0.0005	−0.0002	0.0003
V12, F3	−0.0134	−0.0092	0.0042
V13, F1	−0.0074	−0.0061	0.0013
V13, F2	−0.0056	−0.0012	0.0044
V13, F3	−0.0043	−0.0085	−0.0042
V14, F1	−0.0070	−0.0041	0.0029
V14, F2	−0.0021	−0.0018	0.0003
V14, F3	−0.0027	−0.0061	−0.0034
V15, F1	0.0037	0.0017	−0.0020
V15, F2	−0.0061	−0.0014	0.0047
V15, F3	−0.0056	−0.0073	−0.0017

Only one computational approach to minimizing an EFAC function was developed in this article. While it worked well in our example and simulation, we are hopeful that future research can develop alternative optimization methods that may improve on our work, e.g., in obtaining faster rates of convergence or improved ability to handle difficult situations such as near-linear dependence among factors.

Although we showed that the proposed EFAC methodology can yield results that are essentially equivalent to those from standard methods, our comparisons are illustrative rather than exhaustive. Further research is needed to compare EFAC to a wider variety of factor analytic methods, as well as under a wider variety of simulated conditions and with many more empirical examples. A related important question to consider is whether any variants of this methodology actually can yield improvements over existing methods and, if so, under what conditions. If not, our results will be of interest mainly in providing a new theoretical perspective on the relations between components and factors. Of course, we hope the latter contribution also may imply new research directions, e.g., in the use of (7) to determine the correlations between common factors and components in various limiting conditions such as when $p \to \infty$ with k fixed.

In this article, for simplicity we considered only the exploratory factor analysis model. It is possible to extend our approach to relate components to other latent variable models as well. Similarly, we considered only the simplest possible view of component analysis. There is a huge and continuously growing literature on alternative approaches to components that also might be considered (e.g., Johnstone & Lu 2009; Nadler, 2008; Shen & Huang 2008; Witten, Tibshirani, & Hastie, 2009).

Acknowledgements

Research supported by grants 5K05DA000017-32 and 5P01DA001070-35 from the National Institute on Drug Abuse to P.M. Bentler and grant 4R44CA137841-03 from the National Cancer Institute to P. Mair. Bentler acknowledges a financial interest in EQS and its distributor, Multivariate Software.

References

Bentler, P.M., & Kano, Y. (1990). On the equivalence of factors and components. *Multivariate Behavioral Research*, *25*, 67–74.

de Leeuw, J. (2004). Least squares optimal scaling of partially observed linear systems. In K. van Montfort, J. Oud, & A. Satorra (Eds.), *Recent developments on structural equation models: theory and applications* (pp. 121–134). Dordrecht: Kluwer Academic. See also http://preprints.stat.ucla.edu, Preprint #360.

Guttman, L. (1956). Best possible systematic estimates of communalities. *Psychometrika*, *21*, 273–285.

Harman, H.H. (1976). *Modern factor analysis* (3rd ed. Revised). Chicago: University of Chicago Press.

Jennrich, R.I. (2002). A simple general method for oblique rotation. *Psychometrika*, *67*, 7–20.

Johnstone, I.M., & Lu, A.Y. (2009). On consistency and sparsity for principal components analysis in high dimensions. *Journal of the American Statistical Association*, *104*(486), 682–693.

Mulaik, S.A. (2010). *Foundations of factor analysis* (2nd ed.). Boca Raton: Chapman & Hall/CRC.

Nadler, B. (2008). Finite sample approximation results for principal component analysis: a matrix perturbation approach. *Annals of Statistics*, *36*(6), 2791–2817.

Schönemann, P.H. (1966). A general solution of the orthogonal Procrustes problem. *Psychometrika*, *31*, 1–10.

Shapiro, A. (1984). A note on the consistency of estimators in the analysis of moment structures. *British Journal of Mathematical and Statistical Psychology*, *37*, 84–88.

Shen, H., & Huang, J.Z. (2008). Sparse principal component analysis via regularized low rank matrix approximation. *Journal of Multivariate Analysis*, *99*(6), 1015–1034.

Slutsky, E. (1925). Über stochastische Asymptoten und Grenzwerte. *Metron*, *5*(3), 3–89 (in German)

Witten, D.M., Tibshirani, R., & Hastie, T. (2009). A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. *Biostatistics*, *10*(3), 515–534.

Manuscript Received: 11 JUN 2010
Final Version Received: 24 NOV 2010
Published Online Date: 1 JUN 2011