Multivariate Behavioral Research, 31 (4), 539-550

Dispelling Some Myths About Factor Indeterminacy

James H. Steiger
University of British Columbia

A simple numerical example helps illuminate some of the issues discussed by Maraun (1996), and also helps dispel some of the myths connected with the *posterior moment position*.

Maraun (1996) evaluates, with admirable clarity and succinctness, several conceptual positions regarding the phenomena of factor indeterminacy. He dismisses two of the positions as either irrelevant or misguided, and concentrates his attention on two points of view which he calls the *alternative solution position* (ASP) and the *posterior moment position* (PMP). Maraun argues very convincingly that the *alternative solution position* (ASP) view of indeterminacy is the correct one. According to Maraun, PMP "conflates the meaning of a *latent common factor* with senses of the concept of *factor* that are external to the model. When the criterion for latent common factor is properly explicated, it is clear that there is indeed a fundamental indeterminacy to the model."

In this commentary, I present a simple numerical demonstration, (a) as an aid to those unfamiliar with the intricacies of indeterminacy, and (b) as a challenge to some who *are* familiar, and claim indeterminacy has no important consequences. This numerical example seems to demonstrate, rather convincingly, that some of the assertions of the proponents of PMP are incorrect.

A Signal from Space?

Suppose it is *believed* that a signal may be emitted ten times (*and only ten times*) at one hour intervals from point X, starting at 1:00 P.M. Point X can never be observed directly. Receivers are constructed to measure the signal at points Y_1, Y_2, Y_3, Y_4. However, the signal is "jammed" by a noise countersignal at each receiving point. An additional signal, at point W, is received directly without any noise degradation. Fortunately, it is known from intelligence sources that (a) all signal and noise distributions have zero means, (b) the signal and noise components are additive, and (c) they are precisely uncorrelated over the ten observations taken, and (d) that the ten signals have a variance of exactly 1. Moreover, suppose that the distribution of $\mathbf{Y}$ is discrete multivariate, with only ten 4-tuples assigned positive probability, the ten 4-tuples equally likely, and so the ten observations happen to mirror exactly the population multivariate distribution.

Clearly, the signal believed to exist can be considered a random variable X that satisfies the definition of a common factor for the 4 random variables in the vector $\mathbf{Y}$. If the assumptions (a)-(d) are correct, the common factor model is an appropriate data generation model for these data. Now suppose that the 10 observations on $\mathbf{Y}$ are received at the 4 stations, one each hour, starting at 1:00 P.M. These ten observations are shown in Table 1.

Note that the information in Table 1 for $\mathbf{Y}$ may be read in either of two ways. One can regard it as a finite list of observations, in which case one should transpose the factor analysis equations appropriately to adjust for the fact that the scores are in columns, that is, $\mathbf{Y} = X\boldsymbol{\Lambda}' + \boldsymbol{\delta}$, *or* as the discrete multivariate distribution for random variable $\mathbf{Y}$, with each 4-tuple having probability 1/10, in which case the equations hold as in Maraun's (1996) article.

We have several questions of interest, which we hope the factor analysis model, coupled with our data analysis tools, can shed some light on:

1. How much noise is degrading each signal?
2. What are the values of the signals? In particular, was the signal at 3:00 P.M. greater than the signal at 4:00 P.M.?
3. Do the signals show any intelligent pattern?
4. Is there any correlation between the signal at $\mathbf{X}$ and the signal at W? If so, what is it?

This problem differs from the typical "real world" application of factor analysis in two very important respects.

1. We are assuming for simplicity that the entire population is available, and there is no sampling error at all in estimating the covariance matrix $\boldsymbol{\Sigma}$ for the variables in $\mathbf{Y}$, and
2. We are taking it as a certainty that the assumptions of the factor model actually hold.

Typically, neither of these two assumptions could be taken for granted. Only a small sample from a larger population would be available, and so our problems would be exacerbated by the additional complexities of sampling variability and statistical estimation. Moreover, we would have to decide on the basis of imperfect evidence whether the factor model is appropriate at all. Such problems tend to obscure the issues Maraun (1996) is trying to clarify, so I have eliminated them from consideration.

Table 1
Factor Indeterminacy: A Numerical Example

$$\mathbf{Y} = \begin{bmatrix} 0.905 & 1.641 & 0.203 & -1.041 \\ -0.591 & -0.598 & -0.929 & -0.192 \\ -0.501 & 0.37 & 1.848 & 1.752 \\ -0.488 & -0.495 & 0.74 & -0.402 \\ -0.785 & -1.101 & -0.074 & -0.794 \\ -1.598 & 1.216 & -0.404 & -0.9 \\ 0.749 & 0.514 & -1.703 & 1.084 \\ -0.079 & -0.343 & -0.727 & 1.454 \\ 2.132 & 0.576 & 1.226 & -0.001 \\ 0.255 & -1.779 & -0.182 & -0.96 \end{bmatrix} \quad W = \begin{bmatrix} 0.344 \\ -0.332 \\ 0.805 \\ -2.19 \\ -0.688 \\ 0.309 \\ -0.363 \\ 0.14 \\ 1.898 \\ 0.077 \end{bmatrix} \quad \mathbf{\Lambda} = \begin{bmatrix} 0.5 \\ 0.4 \\ 0.3 \\ 0.2 \end{bmatrix}$$

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & 0.2 & 0.15 & 0.1 \\ 0.2 & 1 & 0.12 & 0.08 \\ 0.15 & 0.12 & 1 & 0.06 \\ 0.1 & 0.08 & 0.06 & 1 \end{bmatrix} \quad \mathbf{\Psi}^2 = \begin{bmatrix} 0.75 & 0 & 0 & 0 \\ 0 & 0.84 & 0 & 0 \\ 0 & 0 & 0.91 & 0 \\ 0 & 0 & 0 & 0.96 \end{bmatrix} \quad p = 0.775$$

$$\mathbf{I}_1 = \begin{bmatrix} 0.398 \\ -0.284 \\ -0.314 \\ 1.949 \\ -0.055 \\ -0.794 \\ 0.608 \\ -0.232 \\ -0.636 \\ -0.64 \end{bmatrix} \quad D_x = \begin{bmatrix} 0.742 \\ -0.616 \\ 0.491 \\ -0.241 \\ -0.743 \\ -0.485 \\ 0.245 \\ -0.092 \\ 1.261 \\ -0.563 \end{bmatrix} \quad X_1 = \begin{bmatrix} 1.14 \\ -0.9 \\ 0.176 \\ 1.708 \\ -0.798 \\ -1.278 \\ 0.853 \\ -0.323 \\ 0.625 \\ -1.203 \end{bmatrix} \quad \mathbf{I}_2 = \begin{bmatrix} 0.258 \\ -0.384 \\ 0.509 \\ -0.759 \\ 1.743 \\ -0.515 \\ 0.755 \\ -0.908 \\ -0.261 \\ -0.437 \end{bmatrix} \quad X_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \quad X_3 = \begin{bmatrix} 0.484 \\ -0.232 \\ 0.518 \\ 0.518 \\ -2.486 \\ 0..031 \\ -0.509 \\ 0.817 \\ 1.522 \\ -0.125 \end{bmatrix}$$

$\text{corr}(X_1, W) = -0.202 \quad \text{corr}(X_2, W) = 0.399$

The ten observations on $\mathbf{Y}$ and W are given in Table 1, along with some summary statistics associated with them. Notice that, since the entire population of observations is available, I have computed the covariance (and correlation) matrix $\mathbf{\Sigma}$ using the population formulae (with N, rather than N - 1 in the denominator).

It takes only a few moments to type $\mathbf{\Sigma}$ into a standard factor analysis program to fit the common factor model to the data. One discovers that a single common factor model fits the data perfectly, with the values for $\mathbf{\Lambda}$ and

$\mathbf{\Psi}^2$ given in Table 1. The factor model fits the data perfectly, $\mathbf{\Psi}^2$ is identified, so we can specify the amount of noise variance degrading each signal. It turns out there is a substantial amount of noise degrading each signal, but that there is more noise degrading signal Y_4 (.96) than is degrading signal Y_1 (.75). Hence, the factor model *can* answer question 1. Note also that, if all the assumptions of the model are correct, a component model will not produce a correct answer to this question.

Subsequent questions in our list require us to decide on the value of the signal emanating from X at the various hours on the basis of the values obtained at Y_1, Y_2, Y_3, Y_4. Unfortunately, in accordance with Maraun's (1996) Equation 3, there are infinitely many sets of random variables X and $\boldsymbol{\delta}$ that might have generated $\mathbf{Y}$ via the factor analysis equation $\mathbf{Y} = \mathbf{\Lambda}X + \boldsymbol{\delta}$. They can be specified via Maraun's Equation 3. For example, D_x can be calculated as in Table 1. This "determinate" part of X is often referred to as "regression factor score estimates" by factor analysis programs. There are infinitely many ways to construct the "indeterminate component" I_x. Since in this case I_x can be vector of finite length, we can construct numbers. Two such examples are shown in Table 1. The reader may verify that each of the two candidates for the common factor X fits the observed data, the factor pattern $\mathbf{\Lambda}$, and all the stipulations of the factor model perfectly. (One may, incidentally, calculate a table of $\boldsymbol{\delta}$ values corresponding to the X values in Table 1 via the equation $\boldsymbol{\delta} = \mathbf{Y} - X\mathbf{\Lambda}'$, for any of the alternative Xs. These will satisfy all the stipulations of the common factor model. For brevity, I have not included them in Table 1.)

Suppose, upon constructing solution X_2 as in Table 1, Person A announces that "intelligent life has been found in a distant galaxy." His justification is that X_2 has a beautiful, orderly structure that suggests it might be an attempt by intelligent life to communicate with us. Later in the day, Person B announces that there need not be intelligent life on the distant galaxy, because X_1 and X_2 are equally valid common factors of $\mathbf{Y}$, and X_1 shows no orderly pattern at all. Neither does D_x.

While X_1 and X_2 have identical covariance relations with $\mathbf{Y}$, they give rather different answers to questions 2, 3, and 4. Specifically, X_1 has a strength of .176 at 3:00 and a much higher value of 1.708 at 4:00. On the other hand, X_2 has a value of 1.00 at 3:00 and a much *lower* value of -1.00 at 4:00. Factor X_2 gives a resounding "yes" in answer to question 3, while X_1 gives an equally resounding "no." These two equally valid factors have a correlation of .399. Note that it is possible to construct factors that correlate as little as -.202 with each other with these data. X_3 is an example of a factor that correlates -.202 with X_2. When equally valid factor can correlate so minimally with each other, it seems questionable to talk of having "found the factor X" via factor analysis. Yet that is precisely what factor analysts often do.

Table 1 also demonstrates that factor indeterminacy can have implications at the covariance level, *and the random variable level*, not just "for factor scores." The different factors have different correlations with variable W, an external variable. For example, X_1 correlates approximately -.2 with W, while X_2 correlates approximately +.4.

On "Understanding" Random Variables

Clearly, then, the possible factors X_1 and X_2 are different. Clearly, it is not meaningful to refer to factor analysis of $\mathbf{Y}$ as having determined "a" single factor for $\mathbf{Y}$. There are *many different* factors, having *different* multivariate distributions with $\mathbf{Y}$, which all fit the factor model with the same $\mathbf{\Lambda}$.

Again, I emphasize that the columns of 10 numbers in Table 1 may be conceptualized as the possible values of discrete random variables having a discrete multivariate distribution, with each of the 10 observations having equal probability. If one takes the ten numbers in X_2 and places them alongside $\mathbf{Y}$, one obtains the possible values in a 5-variate multivariate discrete distribution. Clearly, a different distribution results if you put the numbers in X_1 alongside $\mathbf{Y}$. Yet both versions of X fit the factor model perfectly. Those who claim that proponents of the "distinct solution" do not "understand" the concept of a random variable well enough are simply wrong. It is indeed possible to define "different" *random variables* X, $\boldsymbol{\delta}$ which give rise to $\mathbf{Y}$ through the factor model, with the same $\mathbf{\Lambda}$ and $\mathbf{\Psi}^2$. For example, in Table 2 (next page), I present the discrete multivariate distribution of $\mathbf{Y}$ and X_1. One notes that such a distribution fulfills all the requirements for X to be a factor of $\mathbf{Y}$. If one took the values for X_2 from Table 1 and replaced the X_1 values in Table 2 with these numbers, one would have an alternative *random variable* whose multivariate distribution is fully specified, and which also satisfies all the stipulations for an X. One can say that both X_1 and X_2 satisfy the definition of a factor of $\mathbf{Y}$.

Table 2
The Discrete Multivariate Distribution for Random Variables Y_1, Y_2, Y_3, Y_4, X_1

Y_1	Y_2	Y_3	Y_4	X_1	$\Pr(Y_1, Y_2, Y_3, Y_4, X_1)$
0.905	1.641	0.203	-1.041	1.14	.10
-0.591	-0.598	-0.929	-0.192	-0.9	.10
-0.501	0.37	1.848	1.752	0.176	.10
-0.488	-0.495	0.74	-0.402	1.708	.10
-0.785	-1.101	-0.074	-0.794	-0.798	.10
-1.598	1.216	-0.404	-0.9	-1.278	.10
0.749	0.514	-1.703	1.084	0.853	.10
-0.079	-0.343	-0.727	1.454	-0.323	.10
2.132	0.576	1.226	-0.001	0.625	.10
0.255	-1.779	-0.182	-0.96	-1.203	.10

Bartholomew (1981) preferred to look at this phenomenon somewhat differently. His view emanates from the following simple idea. Suppose you view the role of the factor analyst as determining the distribution of X given $\mathbf{Y}$. To simplify the discussion, assume temporarily that only two alternative solutions X_1 and X_2 are available. Suppose you view the role of the factor analyst as determining the distribution of X given $\mathbf{Y}$. If you observe values of 0.905, 1.641, 0.203, -1.041 on Y_1, Y_2, Y_3, Y_4, you could say, referring to Table 1, that X is either +1.14 or +1.00. But Bartholomew's view is that you should think about the situation differently, from a Bayesian perspective. Assume that there is only one random variable X, and that the conditional distribution of X given $\mathbf{Y}$ is not a point distribution. Rather, it assigns probability .50 to the outcome +1.14 and .50 to the outcome +1.00. Note that, in either case, you end up stating that, given observed values of 0.905, 1.641, 0.203, -1.041 on Y_1, Y_2, Y_3, Y_4, X could be +1.14 and could be +1.00.

The important thing is that, not only is this way of thinking about indeterminacy not particularly helpful (in terms of solving the problem), the belief that it is in some sense the correct way to think about indeterminacy seems based on little more than a personal prejudice that has nothing to do with the mathematics of the problem. To understand this more fully, let's continue to assume for the moment that X_1 and X_2 are the only candidates for X we can think of. Clearly X_1 satisfies *all* the mathematical requirements for a random variable fitting the factor model, and clearly the posterior distribution of X_1 given $\mathbf{Y}$ *is* a point distribution.

We can understand Bartholomew's (1981) idea of a conditional distribution of X, $\boldsymbol{\delta}$ given $\mathbf{Y}$. What remains nebulous is precisely what kind of random variable X he is talking about. Consider the *discrete multivariate probability distribution* in Table 3. Notice that, this is simply a probabilistic mixture where the multivariate distributions of $\mathbf{Y}, X_1$ and $\mathbf{Y}, X_2$ as defined in Table 1 are both given equal probability. Note that the *conditional distribution* of X, $\boldsymbol{\delta}$ given any of the ten possible 4-tuples observable on $\mathbf{Y}$ assigns probability .5 to each of two possible values of X. For example if you observe values of 0.905, 1.641, 0.203, -1.041 on Y_1, Y_2, Y_3, Y_4, you see from Table 3 that only two values of X (+1.14 or +1.00) are possible, and so the conditional distribution of X given that particular observed 4-tuple on $\mathbf{Y}$ assigns probability .5 to those two values. One can quickly verify that the

Table 3
The Discrete Multivariate Distribution for Random Variables Y_1, Y_2, Y_3, Y_4, X_b

Y_1	Y_2	Y_3	Y_4	X_b	$\Pr(Y_1, Y_2, Y_3, Y_4, X_b)$
0.905	1.641	0.203	-1.041	1	.05
-.591	-0.598	-0.929	-0.192	-1	.05
-0.501	0.37	1.848	1.752	1	.05
-0.488	-0.495	0.74	-0.402	-1	.05
-0.785	-1.101	-0.074	-0.794	1	.05
-1.598	1.216	-0.404	-0.9	-1	.05
0.749	0.514	-1.703	1.084	1	.05
-0.079	-0.343	-0.727	1.454	-1	.05
2.132	0.576	1.226	-0.001	1	.05
0.255	-1.779	-0.182	-0.96	-1	.05
0.905	1.641	0.203	-1.041	1.14	.05
-0.591	-0.598	-0.929	-0.192	-0.9	.05
-0.501	0.37	1.848	1.752	0.176	.05
-0.488	-0.495	0.74	-0.402	1.708	.05
-0.785	-1.101	-0.074	-0.794	-0.798	.05
-1.598	1.216	-0.404	-0.9	-1.278	.05
0.749	0.514	-1.703	1.084	0.853	.05
-0.079	-0.343	-0.727	1.454	-0.323	.05
2.132	0.576	1.226	-0.001	0.625	.05
0.255	-1.779	-0.182	-0.96	-1.203	.05

random variable X defined in the discrete multivariate distribution of Table 3 also satisfies all the stipulations of the common factor model. It too is a factor of **Y**! But it is not the same random variable as either X_1 or X_2, and has no more claim to being "correct."

The example of Tables 2 and 3 is of course a simplification. There are infinitely many X_i we could construct, so the distribution of X given **Y** is potentially continuous, if one accepts the infinite analog of X_b. If one assumes equal priors, it assigns equal posterior probability to all possible values of X given **Y**. This can only make things worse, not better.

Behavior Domain Notions and Their Implications

Mulaik and McDonald (1978) proposed an *infinite behavior domain* position. This position has merit, though I fear many have simply used it as an argument for not thinking about indeterminacy, rather than actually exploring its implications. It is based on the idea that, as the number of variables in **Y** goes to infinity, if **Y** continues to fit the 1-factor model, indeterminacy vanishes. The upshot of this is that, if you are sampling your variables from an infinite domain *that fits a single factor model,* there is in the infinite domain really only one determinate X, and it is one of the infinity of possibilities available for X in the finite set of variables you have observed so far. Consider the numerical example in Tables 1-3. If you chose one of the possible factors, say, X_1, and generated several thousand other **Y** variates from the distribution of X_1, you would discover that, if you factor analyzed the *new* **Y** that included all these variates, all the possible candidates for X would now be virtually identical. This follows from the elementary notion that, if you add enough predictors to a multiple regression equation, the multiple correlation eventually converges to 1.

Maraun (1996) dismisses the *infinite behavior domain* (IBD) argument. There are at least two reasons for doing so. First, the IBD argument is not about the factor model of equation 1, but a different model, in which the p observed variables are but a sample from a larger domain. So rather than solving the problem, it simply redefines the model. Second, the IBD argument provides an interesting way of thinking about the status of common factors, but it is empirically empty, unless one actually chooses to test its provisions. Clearly, one can always *assume* there are more variables out there that fit the factors which fit the variables one has now. This is rather like an experimenter "solving" the problem of an overly wide confidence interval for the correlation between two variables by assuming that he/she has sampled 45 observations from a larger domain with the same correlation as the present observations! Indeed, many very difficult problems in statistics could be "solved" with the aid of such an approach. So while the infinite domain *model* adds an interesting new philosophical angle to our ways of thinking about indeterminacy, it has no practical implications unless you should decide to test it.

Testing the infinite domain *model* would require gathering more variables, to somehow test the notion that one is in fact sampling from an infinite behavior domain. Of course factor analysts often have absolutely no intention of doing that, with good reason. First, many have already exhausted their efforts gathering the variables for the first factor analysis. Second, all the evidence seems to suggest that, generally, the number of factors required to obtain an adequate fit to data tends to go up as a relatively constant function of the number of variables. Gathering more variables might prove embarrassing as well as time-consuming, because for indeterminacy to vanish, the ratio of number of factors to number of variables has to go to zero in the limit. It is extremely unlikely for this to occur. Perhaps the following mathematics will make clear why.

Theorem 1

Let p-variate random vector $\mathbf{Y}$ satisfy a single factor model with "factor" X, with factor pattern $\mathbf{\Lambda}$ as in Maraun's (1996) Equation 1. Suppose variable W is added to the current test variables in $\mathbf{Y}$. For simplicity, assume all variables have unit variances, so that $\mathbf{\Sigma}$ is a correlation matrix. A necessary and sufficient condition for W to fit the single common factor model for one of the possible versions of factor X is that $\sigma_{yw} = \mathbf{\Lambda}c$, for some scalar c for which $0 < c^2 < 1$.

Proof

Let $\mathbf{Z}$ be the random vector with $\mathbf{Y}$ augmented by W. The augmented factor pattern $\mathbf{\Lambda_z}$ must be of the form

$$\mathbf{\Lambda_z} = \begin{pmatrix} \mathbf{\Lambda} \\ c \end{pmatrix}.$$

The result then follows directly from Maraun's (1996) Equation 2, since $\mathbf{\Sigma_z} = E(\mathbf{ZZ}') = \mathbf{\Lambda_z \Lambda'_z} + \mathbf{\Psi}^2_{\mathbf{z}}$, and

$$\Lambda_z\Lambda'_z = \begin{pmatrix} \Lambda\Lambda' & \Lambda c \\ \Lambda c & c^2 \end{pmatrix}.$$

■

It is easy to see that such a W must follow a very stringent restriction with respect to the multidimensional space occupied by the random variables in $\mathbf{Y}$. Specifically, the part of W predictable from $\mathbf{Y}$ must be colinear with the original determinate part of X. This means that that part of W predictable from the original $\mathbf{Y}$ must not have any predictability from p - 1 dimensions orthogonal to $D_x = \Lambda'\Sigma^{-1}\mathbf{Y}$. The proof is straightforward.

Theorem 2

A equivalent necessary and sufficient condition to that given for variable W in Theorem 1 is that the part of W linearly predictable from $\mathbf{Y}$ be colinear with the determinate part of X in the original battery, that is, $W = cD_x + W_{\perp \mathbf{Y}}$, where $E(\mathbf{Y}W_{\perp \mathbf{Y}}) = \mathbf{0}$.

Proof

By well known results in multiple regression, W may be written as the sum of two random variables, one of which is linearly predictable from $\mathbf{Y}$, and one of which is orthogonal to $\mathbf{Y}$. From the standard formula for beta weights in linear regression,

$$\begin{aligned} W &= \hat{W}_y + W_{\perp y} \\ &= B'\mathbf{Y} + W_{\perp y} \\ &= \sigma'_{wy}\Sigma^{-1}_y\mathbf{Y} + W_{\perp y} \end{aligned}$$

However, from the result of Theorem 1, this becomes

$$\begin{aligned} W &= c\Lambda'\Sigma^{-1}_y\mathbf{Y} + W_{\perp y} \\ &= cD_x + W_{\perp y} \end{aligned}$$

■

Hence, as the *infinite behavior domain* gets larger and larger, the proportion of "available dimensions" in the test battery that the next new variable can have *any* correlation with gets smaller and smaller. **Y** can have 30 variables, but the next W can only correlate at all with a single linear composite of them. It is not difficult to see that, as the number of variables increases, it is increasingly unlikely for any additional variable to satisfy this restriction, even approximately, in practice. This is why the number of factors tends to increase in published factor analytic studies as the number of variables increases.

In spite of my generally skeptical reaction toward the IBD approach, I see some hope for resolution between the IBD and ASP positions. The seminal paper of Schönemann and Wang (1972) included the suggestion that as a practical response to indeterminacy, the determinacy index (the minimum correlation between equivalent factors) be computed. There is an excellent reason for doing so. If, for example, one were to find that the minimum correlation between alternative candidates for factor X is .997, it would seem that indeterminacy has virtually vanished as a problem. Moreover, if one is actually able to sample enough variables to both fit a low dimensionality factor model *and* achieve a high determinacy index, then it would appear there is empirical support for the notion one is in fact sampling from a behavior domain. This suggests that a way out of the indeterminacy dilemma is to take the more useful aspects of Mulaik and McDonald (1978) seriously, and insist that all "factors" used in practice be based on enough variables and observations so that the (a) the determinacy index is high, and (b) population fit of the factor model can be shown to be good.

Some Conclusions about Implications

By examining Maraun's (1996) summary, carefully considering the numerical example in Tables 1-3, and tracing the history of factor indeterminacy, we can arrive at several conclusions. First, Maraun is correct in his fundamental assertion. The alternative solution position is the correct one, and the posterior moment position is simply incorrect. Second, factor indeterminacy *does* have practical implications. We saw in our numerical example that the correlation between an external variable W and factor X is no more determinate than factor X itself. Several questions definitely worth answering must be left unanswered while factors are indeterminate.

In the final analysis, perhaps the simplest yet most telling messages common to the oft-conflicting work on indeterminacy are that high quality, well planned measurement procedures can alleviate indeterminacy problems. Ironically, such measures are in fact *components*, since the only truly determinate "factor" is one that is perfectly predictable from **Y**. It seems high-reliability linear composites are what the factor analytic community has been looking for all along, and perhaps it is no coincidence that Norman Cliff (1995) devoted his Saul Sells Memorial Lecture to discussion of methods for obtaining such composites. Perhaps it is time factor analysts realized that a "determinate common factor" by another name will still measure as well.

References

Bartholomew, D. (1981). Posterior analysis of the factor model. *British Journal of Mathematical and Statistical Psychology*, *34*, 93-99.

Cliff (1995, October). Weights to improve reliability or reliable component analysis. Paper presented at the annual meeting of the Society of Multivariate Experimental Psychology, Bellingham, WA.

Maraun, M. (1996). Metaphor taken as math: Indeterminacy in the factor analysis model. *Multivariate Behavioral Research*, *31(4)*, 517-538.

Mulaik, S. & McDonald, R. P. (1978). The effect of additional variables on factor indeterminacy in models with a single common factor. *Psychometrika*, *43*, 177-192.

Schönemann, P. H. & Wang, M. M. (1972). Some new results on factor indeterminacy. *Psychometrika*, *37*, 61-91.