5. When a tough job comes up my superior has the technical "know how" to get it done.
6. It is reasonable for my superior to decide what he (she) wants me to do.
7. My superior has specialized training in his (her) field.
8. My superior is justified in expecting cooperation from me in work-related matters.
9. My superior can fire me if my performance is consistently below standards.
10. My superior does *not* have the expert knowledge I need to perform my job.
11. My superior can provide opportunities for my advancement if my work is outstanding.
12. I *don't* want to identify myself with my superior.
13. My superior's position entitles her (him) to expect support of her (his) policies from me.
14. My superior can suspend me if I am habitually late in coming to work.
15. My superior *cannot* get me a pay raise even if I do my job well.
16. My superior can see to it that I get no pay raise if my work is unsatisfactory.
17. I prefer to do what my superior suggests because he (she) has high professional expertise.
18. My superior has considerable professional experience to draw from in helping me to do my work.
19. I admire my superior because she (he) treats every person fairly.
20. My superior can fire me if I neglect my duties.
21. I like the personal qualities of my superior.
22. If I put forth extra effort, my superior can take it into consideration to determine my pay raise.
23. My superior's position does *not* give him (her) the authority to change the procedures of my work.
24. I want to develop a good interpersonal relationship with my superior.
25. My superior is *not* the type of person I enjoy working with.
26. I should do what my superior wants because he (she) is my superior.
27. My superior can get me a bonus for earning a good performance rating.
28. My superior can recommend a promotion for me if my performance is consistently above average.
29. My superior has the right to expect me to carry out her (his) instructions.

Multivariate Behavioral Research, 31 (4), 517-538

Metaphor Taken as Math: Indeterminacy in the Factor Analysis Model

Michael D. Maraun
Simon Fraser University

The issue of indeterminacy in the factor analysis model has been the source of a lengthy and on-going debate. This debate can be seen as featuring two relevant interpretations of indeterminacy. The *alternative solution* position considers the latent common factor to be a random variate whose properties are determined by functional constraints inherent in the model. When the model fits the data, an infinity of random variates are criterially latent common factors to the set of manifest variates analyzed. The *posterior moment* position considers the latent common factor to be a single random entity with a non-point posterior distribution, given the manifest variables. It is argued here that: (a) The issue of indeterminacy centres on the criterion for the claim "*X* is a latent common factor to **Y**"; (b) the alternative solution position is correct, the posterior moment position representing a conflation of the criterion, which is provided by the equations of the model, with metaphors, analogies, and senses of "factor" that are external to the model. A number of implications for applied work involving factor analysis are discussed.

Despite an extensive catalogue of mathematical results, the debate centering on the indeterminacy of the factor analysis model has in no way abated (see e.g., Vittadini, 1989; Mulaik, 1990; Steiger, 1990). The chief reason for this is that what is contentious about indeterminacy is not its mathematical results, but the interpretation of these results, and, most fundamentally, how they rest on the concept of latent variate. The clarification of this issue has been greatly hampered by two facts: (a) There are a number of senses of the concept of *factor* that are often left dangerously undistinguished in the factor analysis literature; and (b) the metaphors and analogies that run through the practice of factor analysis are among the most suggestive in psychometrics, and have always been difficult to distinguish from the mathematics of the model. As Wittgenstein (1953) argued, our forms of representation can mislead. In this article I will attempt to remedy this situation by focussing on the logical issues on which a correct

This research was supported in part by a President's Research Grant awarded to the author by Simon Fraser University.
Correspondence: Dr. Michael Maraun, Dept. of Psychology, Simon Fraser University, Burnaby, B.C., Canada V5A 1S6. Fax: (604) 291-3427.

interpretation of indeterminacy is predicated. I will attempt to clarify what indeterminacy is, and what it is not. It will be suggested that:

1. Indeterminacy is a mathematico-grammatical issue. It centres on the criterion for the claim "X is a *latent common factor* to the n manifest variates $\{Y_1, Y_2, ..., Y_n\}$", and the cardinality of the set of variates that are criterially *latent common factors* to the manifest variates;

2. There have been essentially two relevant interpretations of indeterminacy implicit in the indeterminacy debate. In the present work they are called the *posterior moment position* (PMP) and *alternative solution position* (ASP);

3. Two other positions, the *infinite behavioral domain* argument and what might be called the *scientific usefulness* argument are shown to be external to the issue of indeterminacy;

4. The posterior moment position conflates the meaning of *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 in the model;

5. While external to a clarification of the nature of indeterminacy, it is nonetheless useful to consider implications of the indeterminacy property of factor analysis. A number of implications are examined.

Since the issue of indeterminacy is not affected in any important way by considerations of dimensionality, the simplest case, the single common-factor model, will be considered throughout. In addition, despite the fact that many treatments of indeterminacy have focussed on a purely data analytic version of the factor analysis model, the model considered in this work is the common factor analysis model for n random variates as in Bartholomew (1981).

Mathematics of Indeterminacy

It has been stated previously (e.g., Bartholomew, 1981) that a certain degree of confusion has been brought to the indeterminacy debate by a careless use of terms such as *latent variate*, *factor score*, *factor score estimate*, *regression estimate*, and a corresponding lack of clarity in the specification of the statistical model under consideration. For the record, in the random common factor analysis model the latent factors are random variates with non-point distributions. For these models, estimation is an issue only with respect to the structural parameters (i.e., the loadings and variances of the latent variates) of the model, and not the latent factors themselves. That is, in this context there is no such thing as factor scores, person parameters, or factor score estimates. Instead, prediction (e.g., of the latent variates from the

manifest variates), dependency and correlation are the relevant statistical notions when considering the latent factors themselves. If the latent factors were in fact *scores*, the model under consideration would then be the fixed common-factor score model (Anderson & Rubin, 1956; Bartholomew, 1981; McDonald, 1979; van Der Leeden, 1990). In the fixed-score model, the common factor scores are indeed parameters to be *estimated*, and so what is at issue is the *identification* of these "person parameters", not their indeterminacy.

The 1-dimensional random common factor analysis model for n random variates, $\mathbf{Y}$, is written

(1)
$$\mathbf{Y} = \mathbf{\Lambda} X + \boldsymbol{\delta} \text{ with}$$
$$V(X) = 1,\ C(\boldsymbol{\delta}) = \mathbf{\Psi}^2,\ C(X\boldsymbol{\delta}') = \mathbf{0},\ E(X) = 0, \text{ and}$$
$$E(\mathbf{Y}) = E(\boldsymbol{\delta}) = 0,$$

with the consequence that

(2)
$$\mathbf{\Sigma} = \mathbf{\Lambda}\mathbf{\Lambda}' + \mathbf{\Psi}^2,$$

and in which X represents the latent common factor, $\boldsymbol{\delta}$ is a vector of latent unique factors, $\mathbf{\Psi}^2$ is an $n \times n$ diagonal, positive definite matrix containing the variances of the unique factors, $\mathbf{\Sigma}$, the covariance matrix of $\mathbf{Y}$, is then also nonsingular, $\mathbf{\Lambda}$ is an $n \times 1$ vector of real coefficients called factor loadings, and, without loss of generality, X has unit variance. The defining, and much debated, property of indeterminacy is apparently that when $\mathbf{Y}$ is described by Equation 1, an infinity of sets of random variates possess the required properties to be called latent factors to $\mathbf{Y}$, unless $\mathbf{\Lambda}'\mathbf{\Sigma}^{-1}\mathbf{\Lambda} = 1$, which in practice is never the case. It follows that any admissible set $\{\boldsymbol{\delta}, X\}_i$, must have the following form (McDonald, 1974; Guttman, 1955):

(3)
$$X_i = \mathbf{\Lambda}'\mathbf{\Sigma}^{-1}\mathbf{Y} + pS_i = D_x + I_x \text{ and}$$
$$\boldsymbol{\delta}_i = \mathbf{\Psi}^2\mathbf{\Sigma}^{-1}\mathbf{Y} - \mathbf{\Lambda}pS_i = D_{\boldsymbol{\delta}} + I_{\boldsymbol{\delta}}, \text{ with}$$
$$p^2 = (1 - \mathbf{\Lambda}'\mathbf{\Sigma}^{-1}\mathbf{\Lambda}),$$
$$E(pS_i\mathbf{Y}') = \mathbf{0},\ E(S_i^2) = 1,\ E(S_i) = \mathbf{0}.$$

The components of Equation 3 denoted by D are determinate, in the sense that they are functions of the manifest variables. For any admissible set $\{X, \boldsymbol{\delta}\}_i$, D_X and $D_{\boldsymbol{\delta}}$ are constants. D_X and $D_{\boldsymbol{\delta}}$ turn out to be the linear conditional expectations of X and $\boldsymbol{\delta}$ given $\mathbf{Y}$. The I components, on the other hand, are uncorrelated with the manifest variables, and are arbitrary in the sense that they are not functions of the manifest variables. A different I component is associated with each of the admissible vectors of latent variates. That is, each $\{X, \boldsymbol{\delta}\}_i$ is constructed by choosing a random variate S_i with the properties given in Equation 3. The random variates S_i are therefore somewhat akin to the seeds used in random number generation.

Two Interpretations

The detailed history of indeterminacy provided by Steiger and Schönemann (1978) documents a long and often heated debate centering on the interpretation of the mathematical results of indeterminacy. A number of more recent comments (Mulaik, 1990, 1993, 1994; Steiger, 1990; Vittadini, 1989) suggest that the issue is far from being resolved. Exactly what is to be made of the equations in Expression 3? The contentiousness attending indeterminacy is not really surprising, since it is wedded to an array of fundamental psychometric issues. What is meant by the concept of *latent common factor*? If it is present in the model, what implication does indeterminacy have for the model as a basis for formulating and testing hypotheses about covariance structures? What, if anything, can factor analysis contribute to scientific investigation? What is the logical status of attempts to establish the meaning of a concept via correlations and other empirical results? It is proposed here that the majority, if not all, of the relevant interpretations of indeterminacy are examples of one of two positions. These positions are here referred to as the posterior moment and the alternative solution positions. In this section, an overview of each is provided. It must be noted, however, that there are two additional stances of notable importance. The first is often called the *infinite behavior domain* argument (see, e.g., Mulaik & McDonald, 1978; Williams, 1978). The second is most clearly stated in Mulaik's recent writings on the normative grounding of factor analytic practice (e.g., Mulaik 1993, 1994). This stance might be called the *scientific usefulness* argument. It will be argued that while both contain important insights (for different reasons), they are nevertheless external to the issue of indeterminacy. That is, they do not address questions pertaining to the nature or existence of indeterminacy.

Alternative Solution Position

The alternative solution position is manifest in the work of Wilson (1928), Camp (1932), Guttman (1955), Schönemann and Wang (1972), and

Steiger and Schönemann (1978), among others. The gist of the position can be stated as follows:

$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}^2$ and $\boldsymbol{\Psi}^2$ diagonal, positive definite $\rightarrow \exists\ \infty$ sets of random variates $\{X, \boldsymbol{\delta}\}_i$ such that $\mathbf{Y} = \boldsymbol{\Lambda}X_i + \boldsymbol{\delta}_i$, and

$$C(X_i, \boldsymbol{\delta}_i) = \begin{pmatrix} 1 & \mathbf{0}' \\ \mathbf{0} & \boldsymbol{\Psi}^2 \end{pmatrix}.$$

Each of these sets contains, *by definition*, one common and n unique latent factors for the n variables $\mathbf{Y}$. The latent factors have the following construction (with p as in Equation 3):

$$\begin{pmatrix} X_i \\ \boldsymbol{\delta}_i \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Lambda}'\boldsymbol{\Sigma}^{-1} & p \\ \boldsymbol{\Psi}^2\boldsymbol{\Sigma}^{-1} & -\boldsymbol{\Lambda}p \end{pmatrix} \begin{pmatrix} \mathbf{Y} \\ S_i \end{pmatrix}.$$

What is asserted here can be further unpacked. First, a latent factor is exactly what is specified by Equation 3, and nothing more. Hence, when Equation 1 holds, the factor analysis model provides a criterion that admits an infinity of constructed variates that are *each* latent common factors to $\mathbf{Y}$. That is, when Equation 1 holds, a set S containing admissible random variates (latent factors) is implied, and the cardinality of S is infinity. Second, while the vectors $\{X, \boldsymbol{\delta}\}_i$ are not distinguishable at the level of expectation and variance, they are most certainly so at the level of covariance. Thus one can consider the correlations of the X_i with each other, and with external variables (Steiger, 1979). In particular, the degree of indeterminacy present in a specific factor analytic result is usefully quantified by the minimum correlation between maximally different admissible factors, ρ^* (Guttman, 1955). This index quantifies the maximum dissimilarity of the elements X_i in S. In fact, $2(1 - \rho^*)$ is the squared euclidean distance between the two most dissimilar latent common factors in S. Third, the practice of "interpreting the factor" by examining $\boldsymbol{\Lambda}$ is usually meaningless since $\boldsymbol{\Lambda}$ describes the relationship between $\mathbf{Y}$ and each of an infinity of random variates X_i, some of the X_i possibly having small or even negative correlation (Guttman, 1955). This is but a particular instance of the questionable practice of attempting definition via correlational links (Guttman, 1977). Fourth, the grammar of factor analytic theory is misleading since it consistently implies that there is but one variate that can legitimately be called a latent common factor to **Y**. Finally, the factor analysis model fails according to its stated aim (Garnett, 1919; Spearman, 1933) since it was supposed to be the case that if the model held, that is, if R was rendered conditionally diagonal, only one such variable (*the* factor) was responsible for this diagonalization.

Posterior Moment Position

McDonald's (1974) article was perhaps the first focused attempt to defend the factor analysis model against the alternative solution argument and its implications. Although McDonald's article contains many different ideas on the matter, one is based on what is called here the posterior moment argument. This can be seen in his championing of the squared multiple correlation as the appropriate index of the degree of indeterminacy present in a factor analytic result (McDonald, 1974, p.214-215). Bartholomew (1981) provides a more direct defence of the factor analysis model using this argument, while Holland (1990) gives a characterization of the latent variates of random IRT models in terms of the posterior moment position. The posterior moment position can be stated as follows:

$\boldsymbol{\Sigma} = \boldsymbol{\Lambda\Lambda}' + \boldsymbol{\Psi}^2$ and $\boldsymbol{\Psi}^2$ diagonal, positive definite $\rightarrow \exists$ but one vector of random latent variates $[X, \boldsymbol{\delta}]$ such that $\mathbf{Y} = \boldsymbol{\Lambda}X + \boldsymbol{\delta}$ and

$$C(X, \boldsymbol{\delta}) = \begin{pmatrix} 1 & \mathbf{0}' \\ \mathbf{0} & \boldsymbol{\Psi}^2 \end{pmatrix},$$

but X does not have a point distribution after **Y** has been observed (i.e., it does not have a posterior point distribution).

Several of the important features of this interpretation are listed. First, the "solutions" given by Equation 3 represent a mischaracterization of the issue. The model does not imply *alternative* latent factors. Second, if $\boldsymbol{\Sigma} = \boldsymbol{\Lambda\Lambda}' + \boldsymbol{\Psi}^2$, then by specifying the distribution of $[\mathbf{Y}, X]$, $E(X|\mathbf{Y})$ and $V(X|\mathbf{Y})$ can be derived. As Bartholomew (1981) states, this "posterior analysis" begins at the point at which $\mathbf{Y} = y$ has been observed. In other words, $(X|\mathbf{Y})$ is a random quantity, and knowledge of the properties of the latent common factor X rests on the distribution of $(X|\mathbf{Y})$. One may express one's knowledge of X in terms of posterior moments. The conditional location and scaling of X, $E(X|\mathbf{Y})$ and $V(X|\mathbf{Y})$, each functions of the parameters of the model, are particularly relevant (Bartholomew, 1981, p. 95). The former quantity is, in the multi-normal case, what is known as the regression estimate (McDonald & Burr, 1967). Third, it follows that the issue of the relationship between **Y** and X, and more importantly, the logical status of X, is, as Bartholomew (1981, p. 94) states, "a simple and routine application of elementary probability theory". Fourth, indeterminacy, meaning simply that $(X|\mathbf{Y})$ does not have a point distribution, is actually a simple question of statistical *determination*. That is, just as in any problem of mean-square prediction, the goal is to optimally predict values of X given values of **Y**. However, perfect prediction is rarely achieved (McDonald, 1974, p. 214) and so the investigator must resign himself to the obtainment of incomplete knowledge of X. This is a modern incarnation of the linear unpredictability argument of Spearman (1927). Fifth, since there is but one common factor, the minimum correlation between alternative factors, ρ^*, is of questionable logical standing (e.g., McDonald, 1974). Instead, ρ^2, the standard statistical coefficient of determination, is the appropriate measure to quantify the degree of "knowledge" one has of X.

Metaphor Taken as Math

It would seem then that the alternative solution argument and the posterior moment argument are fundamentally opposed interpretations of the mathematics of indeterminacy, and, at a deeper level, the nature of latent variates. However, the issue of indeterminacy is a mathetico-grammatical issue. It centres on the criterion for the claim "X is a latent common factor to **Y**", or, in other words, on the standard of correctness for this claim. The basic question of indeterminacy, whether there is indeed a multiplicity of latent common factors to **Y** when **Y** is described by the model, requires an explication of this criterion. This is the key to the indeterminacy debate (see Mulaik, 1993, 1994, for a similar interpretation of indeterminacy). The question is then, what are the contents of the right bracket in the following?

(4) X is a latent common factor to $\mathbf{Y} \leftrightarrow \{........\}$

In the first place, the concept of latent variate denotes a random variate for which there exists no criterion of identity external to a latent variable model. That is, the justification for the claim "X is a latent variate" is that the criterion of identity for X is provided by the mathematical equations of a latent variable model itself. This is in fundamental contrast to a manifest variate, which does indeed have a criterion of identity external to the model (this being a precondition for the study of, the phrasing of hypotheses about, etc., such a variate). In fact, in the social sciences manifest variates are commonly treated (rightly or wrongly) as if they are (criterially) measurements of some common language property ϕ (e.g., *weight*, *height*, *self-esteem*, etc.) of individuals. For example, if three sets of numbers were to be factor analyzed, the justification for the claim that the first set were measurements of *height* would rest on their having been recorded or taken according to the rules for the measurement of height, these rules a part of common language. The criterion for the claim is external to the model. This is what makes it a manifest variate. In fact, one could not carry out a factor analysis of *height* and two other variables unless this was the case. What makes a variate latent is *not* then that it is unobservable, hypothesized, possibly causal, "underlying", etcetera, but that its character is determined solely by the equations of a latent variable model. Something that is unobservable, hypothesized, underlying, or possibly causal is unobserved, hypothesized, underlying, or possibly causal, not latent. More will be said on this at a later point.

The point then is that the criterion for the phrase "X is a latent common factor to $\mathbf{Y}$" is manifest in the functional constraints (and side-conditions) given by the equations of the factor analysis model when it describes $\mathbf{Y}$. This is what makes the criterion internal to the model. These constraints are embodied in a rule for the construction of $\{X, \boldsymbol{\delta}\}_i$, and this rule is as given in Equation 3. A latent common factor to $\mathbf{Y}$ is therefore merely a random variate constructed in accord with construction rule Equation 3. A rule is a standard of correctness that governs a practice (Hacker, 1988). So to put it differently, the justification for the claim "X is a latent common factor to $\mathbf{Y}$" is that X was constructed in accord with Equation 3. Hence, Equation 3 is the contents of the bracket in Equation 4, and because Equation 3 admits an infinity of different S_i, an infinity of latent common factors can be constructed to be *criterially* latent common factors to $\mathbf{Y}$. Therefore, the random common factor analysis model is indeed indeterminate. Furthermore, this is not an empirical claim, or property, but a grammatical certainty, following from the rules (represented by the model) that establish what it is to be criterially a latent common factor to $\mathbf{Y}$.

It is not therefore that there is just one latent common factor to $\mathbf{Y}$ and this factor has a non-point posterior distribution. The symbol X in Equation 1 is merely a place-holder for a concept with a specific criterion of application. Instead, one can construct an infinity of random variates that are each criterially a latent common factor to $\mathbf{Y}$. The logical issues inherent in this indeterminacy feature of factor analysis arise also in principal component analysis (PCA). It is therefore fruitful to briefly paint the PCA model in the

colours of indeterminacy. In PCA, a component variate is a random variate whose criterion of identity is given by the component model itself. Hence, pace Mulaik (1993, p. 25), a component is, properly speaking, a latent variate. It has no criterion of identity external to the model. The mere fact that a component is a linear combination of manifest variates is irrelevant to the concept of latency since, as Mulaik himself implies, this fact does not in any way speak to the *criterion* for principal component-hood. Applying the concept of "observable" to the component variate in an attempt to gain cleavage between the two models constitutes merely a questionable usage of the concept of "observable" (see p. 528). The criterion for the claim "C is a principal component to the manifest variates $\mathbf{Y}$" is, as in factor analysis, manifest in a construction rule which embodies a specific set of constraints on a random variate that is criterially a component to $\mathbf{Y}$: $C = \mathbf{v}'(\mathbf{Y} - \boldsymbol{\mu})$, $\mathbf{v}'\mathbf{v} = 1$, in which $\mathbf{v}$ is an eigenvector of $C(\mathbf{Y})$. Note that, as with factor analysis, one does not have to proceed with the construction of the component variate itself (ones interest could be just in the eigenvector $\mathbf{v}$). However, if one did proceed, one would construct the component variate in accord with the equation above. Construction according to *this* equation is what yields a variate that is criterially a component to $\mathbf{Y}$. In factor analysis there is:

$$X_i = \boldsymbol{\Lambda}'\boldsymbol{\Sigma}^{-1}\mathbf{Y} + pS_i \text{ with}$$
$$p^2 = (1 - \boldsymbol{\Lambda}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda}),$$
$$E(pS_i\mathbf{Y}') = \mathbf{0}, E(S_i^2) = 1, E(S_i) = \mathbf{0}.$$

In principal component analysis there is:

$$C = \mathbf{v}'(\mathbf{Y} - \boldsymbol{\mu}).$$

If the issue is the criterion for the claim "X is a latent common factor (component) to $\mathbf{Y}$", then the difference between the two models is nothing more startling than the latitude allowed by the construction rule inherent in each. The fact that factor analysis has always had attributed to it powers far beyond that of PCA speaks more to the powers of the metaphors and analogies that attend the use of the model, than to special powers inherent in the model itself. More will be said on this issue later.

The posterior moment position is founded on the assumption that there is but one latent common factor to $\mathbf{Y}$. But there is no support for this claim. Since what is in question is the criterion for the claim "X is a *latent common factor* to $\mathbf{Y}$", the identity relation of Equation 4 must lead to this conclusion, and yet it does not. It cannot be decided by fiat that there is only one latent common factor to **Y**, any more than can one's desire that there be only one correct answer to $X + Y = 10$ stands as proof that this is the case. A determination of the number of correct answers to $X + Y = 10$ follows from a consideration of the mathematical rule itself. For there to be a justification for the claim that there is a single, particular random variate called *the* latent common factor to **Y**, more would be needed in the brackets of Equation 4. This is roughly Guttman's (1955) argument when he suggests that one would have to leave the context of the common factor analysis model to rid oneself of indeterminacy. It is therefore concluded that the criterion for "latent common factor to **Y**", *as given by the model*, contradicts the statement that there can be but one latent common factor to **Y**, when **Y** is described by the model. This claim, which is the nexus of the posterior moment argument, is at odds with the model itself.

What is not clear at all is *why*, in the posterior moment argument, it is taken as a given that there can be only one latent common factor to **Y**. Both McDonald, at least in his early writing (e.g., 1974), and Bartholomew (1981), among others, hold fast to this tenet, and the singular grammar that accompanies it. Why then is this assumption made when the issue is a purely definitional/mathematical one? While not necessary for a correct characterization of indeterminacy, it is important to consider this question because it speaks to the difficult task of keeping straight what a model, represented symbolically, can and cannot deliver. To paraphrase Wittgenstein (1953), the grammars of both language and math can lead us astray. I believe that the answer to the question is that the posterior moment position is an example of a conflation that is endemic to the practice of factor analysis: The conflation of the criterion of "latent common factor to **Y**", which is manifest in the factor analysis model itself, with metaphors and senses of *factor* that are external to the model (but which are very much a part of the practice of factor analysis).

The factor analysis literature is rife with metaphors, analogies, and distinct notions of *factor*. They are some of the most interesting and suggestive in psychometrics. There are "hypothesized causal factors", "unobserved factors", "principal underlying variates", "underlying attitudes", "mental factors", "factors of mind", "pure measures", "*g*, the genetic basis of intelligence", and "primary mental abilities", each a distinct sense of the concept of *factor*. There is the "error plus truth" analogy drawn from physical measurement. Manifest variates metaphorically "pick-up", "detect", and "tap" metaphorical "underlying factors". The Spearman hypothesis itself, which led to the formulation of the model, was run through

with metaphor. Each of a number of tests $\mathbf{Y}_g$ was viewed as "decomposed" into two components, a single, unitary common factor that was measured by all of the $\mathbf{Y}_g$, and a unique factor that was specific to each of the $\mathbf{Y}_g$. The common factor was viewed as "responsible" for the correlations among the tests in the sense that it was the only thing they "shared". It was the "signal" from the empirical that was "picked up", or "tapped" by the tests. In contrast to all of this, there is the technical concept of "latent common factor to **Y**" as it appears in factor analysis. A latent common factor to **Y** is a random variate constructed according to the construction rule of Equation 3 when the model describes **Y**. It is useful to briefly recall *why* this is the case (i.e., why the model deals with latent variates at all). The Spearman hypothesis (1927) posited the existence of a factor that explains the correlations among a set of tests in the sense that the covariance matrix of the tests, conditional on this factor, was diagonal. Usually a question like this, an empirical question of existence, would require the screening of candidate factors to see which had the desired property of factor-hood. However, factor analysis, Spearman's attempt to phrase his hypothesis in mathematical terms, did not involve an attempt to find such a variate. Instead, the attempt was made to find a diagonal matrix $\boldsymbol{\Psi}^2$ such that $\text{RANK}(\boldsymbol{\Sigma} - \boldsymbol{\Psi}^2) = 1$. If such a matrix is found, then $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}^2$, and $\mathbf{Y} = \boldsymbol{\Lambda}X + \boldsymbol{\delta}$ for at least one set of variates $[X, \boldsymbol{\delta}]$. However, as a result of this particular formulation of the problem, there is no criterion for the statement "X is a common factor to **Y**" external to the model. This is what makes X a *latent* common factor.

It is true, of course, that "fruitful analogies are the go-cart of creativity" (Hacker, 1987), and the same might be said of figurative language in general. However, the figurative language, and various senses of factor that surround the practice of factor analysis must be distinguished from the technical sense of latent common factor inherent in the model. Whereas in the various substantive problems of science there are various notions of the "factor responsible for....", in the model there is the particular, technical concept of the latent common factor. Steiger's (1990) example of a signal that is recorded at several locations provides a good forum to illustrate these points. He considers a scenario in which "one is recording a signal at several locations each subject to independent random noise interference.." (Steiger, 1990, p. 42). Assume that the recordings were factor analyzed, and were in fact described adequately by the model. While the *signal* of this scenario may be a *factor* in some sense, it is certainly *not* a latent common factor (compare with the factors of analysis of variance). To support the initial claim that one *knows* that one is recording *the* signal ϕ one must have relevant evidence, and relevant evidence presupposes a criterion of identity for ϕ external to any model one might consider: *That* is the signal. On the other hand, until one fits the factor model to the data, one *has* no criterion of identity for "*X*, a latent common factor to **Y**". Hence, while the *signal* may be "unobservable", "underlying", etcetera, it is not a latent common factor: it is *the* signal. Is the signal "unmeasurable"? It depends on what one means by the claim. Paradigmatically, measurements are taken of the property "β" of ϕ (Baker & Hacker, 1980). Hence, one might measure the "strength" or "clarity" of the signal. Recordings of the signal, however, are not properly measurements, but instead are *reproductions*. A reproduction has greater affinities to a model, than to a measurement. Is the signal unobservable? Unobservable might mean that the existence of something is in question, or merely that it is something that cannot be seen. In both cases, a criterion of identity is presupposed if these claims are to be anything but vacuous. For it is vacuous to make the claim that *it* is unobserved unless there is a standard of correctness that can be applied to judge the correctness of the claim: for example, what could be made of the claim that "the *grek* is the unobservable entity responsible for...." Is the signal "underlying"? It might make sense to speak of the signal in this way, but this would clearly depend on what one meant by the notion of underlying. Certainly the claim should not suggest a turning over of the recordings to see if the signal is underneath them. In fact, in this context the notion of "underlying" is metaphorical, much like speaking of the underlying causes of the breakup of a country. The point is that when we hypothesize about, theorize about and study, suggested and actual "factors responsible for..." (e.g., a particular signal), *with criteria of application external to a latent variable model*, we are not speaking of latent common factors (regardless of whether various convoluted senses of unobservable, underlying, etc., are invoked). We may, on the other hand, be speaking of *the* signal, socio-economic status, the wind, or the sasquatch in the woods. Likewise it is incorrect to think of the model as a test of whether *the* signal is a factor, in any sense, to **Y**. The factor analysis model says nothing about *the* signal. In fact, if it was known that *the* signal was being recorded at several locations, then the problem would not call for the factor analysis model at all. One could conceivably just compute the conditional covariances of the recordings given the signal.

The scenario described above approaches what is probably the most common external sense of factor, and the one most typically confused with the concept of latent common factor. The idea is this: "we know (or hypothesize) that *it* is there. *It* is unobservable and determines (perhaps causally) the manifest variates in some sense. When the factor analysis model describes **Y**, this is evidence for the correctness of our hunch". A great deal of care must be taken with this sentiment. The claim is certainly coherent, but a careful consideration must be given to the grounds for its justification. The claim is roughly that the factor analysis model plays the role of a litmus test for the presence of *it*, the unobservable determining factor. How might one support such a claim? Assume that it is claimed that *T* is a litmus test for the presence of τ. First, a criterion of identity for τ would be required. The attempt to establish the presence, let alone existence of τ, let alone reaction of *T* in the presence of τ, presupposes a criterion of identity for *it* (i.e., rules that establish what τ denotes). Something cannot be both a criterion of identity for τ and a litmus test for τ. Second, it must be shown that when τ is present, *T* responds in a predictable fashion. Thus, the support for the claim is empirical, and typically probabilistic, *but predicated on a conceptual criterion of identity for* τ. Existence/presence and meaning are logically independent. Is factor analysis a litmus test for the presence of an unobservable *it* that determines the manifest variates in some sense? No it is not. First, the *it* in question could not be a latent common factor to **Y**, because the model itself provides the criterion of identity for an *it* that is a "latent common factor to **Y**". But something cannot be both a criterion of identity for an *it* that is a "latent common factor to **Y**" and a litmus test for the presence of a "latent common factor to **Y**". Second, there has never been any evidence provided that factor analysis acts as a litmus test for the presence/existence of any other *it*. The model was not formulated to be an empirical reflection of the "impact" of an $(n + 1)^{\text{th}}$ variate on n manifest variates. Other multivariate techniques could potentially play this role (consider, for example, the partial correlation of two variates given a third). The factor analysis model is *not* a litmus test of whether *it* is the factor, any more than PCA is a litmus test of whether *it* is the component. Instead, it provides a *criterion* for the claim "*X* is a latent common factor to **Y**". When one analyzes data with factor analysis, what one has are manifest variates and constructed variates called latent factors. Once again, when external senses of *factor* are properly distinguished from the concept of *latent common factor*, the logical unity of PCA and factor analysis is clear.

The primary aim of this article was to clarify the nature of indeterminacy, and assess whether it is indeed a feature of the the random common factor analysis model. Indeterminacy is a mathematico-grammatical issue. It is a consequence of the fact that when the model describes a **Y**, an infinity of random variates can be constructed that are criterially latent common factors to **Y**. The factor analysis model is thus fundamentally indeterminate. Furthermore, there is no *solution* to indeterminacy because it is a grammatical property of the model: that is, it is a result of the very formulation of the model. There are, however, *responses* to indeterminacy, and *implications* of indeterminacy. The next two sections deal with these issues.

Responses to Indeterminacy

It was previously claimed that two positions, the infinite behaviour domain argument (e.g., Mulaik & McDonald, 1978; Williams, 1978) and the scientific usefulness argument (e.g., Mulaik 1993, 1994), were external to the indeterminacy issue. The justification for this claim is that both positions are actually *responses* to the fact of indeterminacy, not interpretations of its nature or existence. This case will now be supported.

Scientific Usefulness Argument

What is called here the scientific usefulness argument is prominent in Mulaik's recent work, which properly characterizes multivariate analysis as a normative practice. On the issue of indeterminacy he states:

> We must break the indeterminate impasse of latent variables by assigning a use for these latent variables and proceed to study the consequences and implications of that use in additional studies, using the initial indicators of our factors as criteria for them (Mulaik, 1994, p. 234).

> One way of dealing with factor indeterminacy has been to argue that what is mathematically possible in the way of numerous alternative constructions for a factor may not apply to the same degree when we seek to embed the variables studied in the larger frameworks of well-established schemas for representing what is in the world (Mulaik, 1993, p. 31).

The idea is that in the face of indeterminacy one proceeds by allowing further investigations to place restrictions on which of the latent common factors are considered to be of interest, or relevant to one's purposes, or "empirically meaningful" (Mulaik & McDonald, 1978). For example, one might only consider those latent common factors that are *useful* in some further scientific context. While an interesting idea, no comment will be made about whether such a "next-step" in the factor analysis program is possible or useful. For the issue of indeterminacy, the only relevant point is that the program envisioned by the scientific usefulness position in no way *eliminates* the indeterminacy property of the model. Indeterminacy is a

grammatical fact of the model. It centres on the criterion for the claim that "X is a latent common factor to **Y**". The considerations put forward by the scientific usefulness position are clearly external to this criterion. In fact, to say that *this* latent common factor has a privileged status because "...", obviously presupposes the criterion itself. All variates constructed in accord with Equation 3 are criterially latent common factors, even if, for certain aims and purposes, some are superior to others. External considerations like scientific usefulness are no more relevant to the criterion for "latent common factor to **Y**" than they would be to the criterion for "component variate to **Y**".

Infinite Behaviour Domain Argument

The infinite behaviour domain position considers the set of n manifest variates analyzed to be a sample from an infinite domain of variates that could have been sampled instead. In the infinite domain, either there are multiple determinate factors, a single unique factor, or no solution. This argument is often taken as overcoming the difficulties implied by the ASP. But this is not the case, for the infinite behaviour domain argument is really an argument for a different model: One in which items are a random facet (see Steiger, 1990, for a similar criticism of this "solution"). As Williams (1978, p.305) states, "..no adequate model has ever been set out before, only finite-dimensional factor analysis equations have been studied". Considerations of the appropriateness of the suggested model aside, it is obvious that switching from model A to model B hardly constitutes a solution to the problems encountered in model A. This strategy is, at best, a response to such problems.

Implications of Indeterminacy

A consideration of the implications of the indeterminacy property of factor analysis is not central to the aims of this article. Yet there are a number of consequences of indeterminacy that are of interest in their own right. These are now described.

The Latent Common Factors are Known

The latent common factors of random common factor analysis are "known". By this it is meant that there is a known rule for their construction, this being rule Equation 3. Terms like hypothetical, unobservable, unmeasurable, and underlying are not properly applied to latent common factors in any non-figurative sense. What makes something latent is that its criterion of identity is given internal to a latent variable model. This brings to the fore a distinction of questionable standing that is often made between components and latent common factors. It is said that components are "observable" because they are functions of the manifest variates. But in what sense is a "synthetic" variate observable (unless one means, circularly, that it is a function of the manifest variates). Is a mean observable? If, by observable, one means the property of being able to take realizations of the variate in question, then a latent common factor is also observable. The application of the property of observability to either components or latent common factors is not strictly correct, and therefore it is not a basis for distinguishing between them. Instead, I think that what is being touched on is an entirely different distinction: The distinction between a concept with a criterion of application internal to a model (i.e., a latent variate), and a concept with a criterion of application that is part of common language. But this distinction suggests the fundamental logical *unity* of components and latent common factors. The difficulty that motivates the misuse of "observable" is that even when one understands the criterion for "component to **Y**" or "latent common factor to **Y**", one is not employing a concept with a common use in the language. One's interest was in *intelligence*, *dominance*, *anxiety*, etcetera, and what one has is a technical concept, a latent common factor or component, with no grammar (i.e., rules) to translate from the latter to the former. This is the problem that is incorrectly attributed to a lack of observability.

What is Estimated is no Mystery

The concept of estimation in the factor analysis model has a rather dubious legacy of attendant confusions. In the case of the random common factor analysis model, factor variates can be predicted, but not estimated, just because they are random variates and not parameters. Thus, the regression "estimate" of each latent common factor when the model describes **Y** is actually a predictor. More specifically, the regression "estimate" is the linear posterior expectation of each latent common factor given **Y**. It has been claimed by a number of psychometricians that, given the indeterminacy of the model, the logical status of these predictions is cloudy. Guttman (1955) has, for example, stated that it is somewhat of a mystery as to what is being predicted. This position, however, is contrary to his own findings. There is no lack of clarity because the latent common factors are those variates constructed according to Equation 3. Since these variates are random, it is not *logically* incorrect to speak of predicting each of them. The question is

why would there be a need to predict an X_i when realizations could be taken on it? Since X_i is just a random variate constructed as in Equation 3, realizations are obtained directly. There of course is an infinity of such latent common factors on which realizations could be taken. This fact, however, cannot be eradicated by focussing on what turns out to be merely the average of the infinity of variates that are criterially latent common factors to **Y** (Steiger & Schönemann, 1978).

The Interpretation of Common Factors

It is a standard practice in the application of the factor analysis model to "interpret" *the* latent common factor by means of an examination of the elements of $\mathbf{\Lambda}$. But exactly what does this mean? What is being sought when one asks for an "interpretation of the music", "an interpretation of the car", an "interpretation of the mean"? I suggest the issue of interpretation, as much as anything, highlights the difference between a latent variate and the manifest variates that are the basis of a factor analytic investigation. The difference does not centre on observability, but on the fact that while we are interested in studying the referents of the common concepts of language (e.g., *intelligence*, *anxiety*, etc.), factor analysis gives us the technical concept of latent common factor, and no rules of translation to bridge the gap. Hence, interpretation, in the context of factor analysis, is the practice whereby one attempts a translation from one set of rules (the rules, manifest in a latent variable model, that establish what a "latent common factor to **Y**" is) into another (the rules for the application of the concepts of language). There is no normative system of translation, and so the investigator must "interpret". If this is the case, then there are a number of problems. First, since when **Y** is described by the model one can construct an infinity of different random variates each which is a latent common factor to **Y**, it is misleading to speak of interpreting *the* factor (Guttman, 1955). Second, it is incorrect to take the criterion for "*X* is a latent common factor to **Y**" to be $\mathbf{\Lambda}$ (which is what one is doing in basing the "interpretation" on $\mathbf{\Lambda}$). This is just one more example of the dubious practice of attempting to establish meaning via correlation. Third, the practice of interpretation invites a straying into the dangerous waters of suggesting that the model is a litmus test for the presence of whatever common notion one has attached to the (technical) latent common factors (e.g., my interpretation of *the* factor is that it is a kind of general intelligence, so therefore the model is *tapping* general intelligence).

On Science

One of the most interesting responses to the mathematical results of indeterminacy is the equating of indeterminacy with the character and difficulties of scientific investigation itself. I will consider two examples that often arise in the factor analysis literature.

Indeterminacy = The Pervasive Indeterminacy of Science

The idea here is that indeterminacy is just an example of the "pervasive indeterminacy that exists throughout science" (Mulaik, 1990, p.54). We are told that it is unreasonable to insist that "scientific concepts are not and must not be indeterminate with respect to the evidence on which they rest" when "...it is now recognized by most philosophers of science (Garrison,1986) that scientific concepts are underdetermined with respect to experience" (Mulaik, 1994, p. 231). This, however, is altogether a false comparison, and is surprising given that it goes against Mulaik's own emphasis on the central place of grammatical rules in the characterization of indeterminacy. It is a category error in which considerations relevant to empirical investigation are wrongly imported to characterize a conceptual issue. In the first place, concepts are not in any way *determined* by experience (Wittgenstein, 1953). Concepts instead have rules of correct application, and a concept's meaning is precisely the set of rules that specify its correct application (Baker & Hacker, 1980). Rules, however, are not right or wrong at all. They are not theories or hypotheses. They are not determined or caused. Instead, they are standards of correctness, and are *constitutive* for experience (Ter Hark, 1990; Wittgenstein, 1953). They are laid down by people, and may be *formulated*, depending on the domain of application, in a number of different ways. They *establish* what there is to theorize about, what there is to determine, what there is to experience, etcetera. On the other hand, it is at least coherent to speak of theories, interpretations, etcetera, as being "determined" by experience, or by empirical phenomena. Indeterminacy, however, centres on a rule, or standard of correctness, for the claim "*X* is a latent common factor to **Y**", a rule formulated in terms of mathematical signs (i.e., Equation 3). Hence, it has nothing to do with empirical determination. It is a *result* of the latitude in admissible variates allowed by a criterion, and nothing more.

The extent of the category error becomes more apparent when an analogy is drawn between indeterminacy and the fact that in science there are rival interpretations and theories:

> Controversies can arise between rival models that account for the same data equally well...In the case of most factor analytic studies, if one were to circulate the tables of factor loadings independently to many individuals and request interpretations for the factors, he would undoubtedly get more than one distinct interpretation for each factor. Some of these interpretations would be nonsensical because they would not fit the knowledge already on hand as to what could produce the correlations among the variables (Mulaik, 1976, p. 253)

Once again, however, the fact that one may construct an infinity of rando variates that are each criterially a "latent common factor to **Y**" does not refle the existence of rival models or theories. The analogy of factors to models inappropriate. A model is a representation *of* something. Different models ca be compared and ranked on a number of relevant criteria, including the quality as representations of theory and their power to suggest useful scientif leads. On the other hand, a latent common factor to **Y** is not a *model* anything. It is a latent common factor to **Y**. Finally, if one did go ahead ar choose a particular X_i, one would not be choosing a particular *interpretation* *the* latent common factor, but a particular latent common factor.

Indeterminacy = The Opportunity for Greater Generalization

In addition, indeterminacy, it is said, gives license to the investigator make broader generalizations:

> ...trying to confine the basis for making a generalization from experience to specific, determinate phenomena already observed and defined, ... by urging the use of specific linear combinations of a set of observed variables (component factors) to stand for what is common to them, may actually get in the way of formulating creative generalization and syntheses that go beyond what is already known or observed... (Mulaik, 1990, p. 54).

On this account, indeterminacy in the factor analysis model is a virtue, th foundation of a statistical tool that squares with the spirit of science itself: "Th implication of this discussion is that determinate models like componer analysis may not have as many scientific virtues as do indeterminate mode like common factor analysis..." (Mulaik, 1990, p. 54). This is a play on the, b now old, urge to impute to the factor analysis model properties that go beyon that of PCA: The factor analysis model somehow goes beyond the informatio present in the manifest variates (For an earlier criticism of this wrong mov see Steiger & Schönemann, 1978). The latent common factors, however, are linear combination of the manifest variates and an arbitrary componen Hence, the only feature of a latent common factor that goes beyond what "known" through the manifest variates is arbitrary. The confusion here seems to result once again from a conflation of metaphors and external senses of *factor* with what the model actually delivers. The factor analysis model is not a litmus test for the presence of underlying causal mechanisms at all. It answers to a much more modest aim: To determine whether a random variate can be constructed that has a particular set of relationships to the manifest variates.

Dimensionality: The Great Trade-off

The definition of unidimensionality inherent in factor analysis is based on Spearman's explanation of why intelligence tests were positively correlated. As mentioned previously, his ideas were phrased mathematically as Equation 1, which leads to the search for a diagonal matrix Ψ^2 such that RANK(Σ - Ψ^2) = 1. This indirect analysis results in the criterion for "X is a latent variate to $\mathbf{Y}$" being formulated as the construction rule of Equation 3. As such, there are *two* senses of unidimensionality inherent in the model. The first, and most well known, sense is simply that RANK(Σ - Ψ^2) = 1. The second is a replacement argument of the type that underlies Guttman's (1977) perfect scale and PCA. A replacement characterization of unidimensionality involves asking whether a single variate can be constructed to replace the set of manifest variates $\{Y_1, Y_2, .., Y_n\}$ according to some criterion. In the case of factor analysis, the question is whether a variate X can be constructed that replaces the individual $\mathbf{Y}_g$ in the sense that $C(\mathbf{Y}|X) = \Psi^2$, with Ψ^2 a diagonal matrix. However, if the factor analysis model describes $\mathbf{Y}$, there is not just one, but an infinity of X_i each of which replaces the $\mathbf{Y}_g$ in exactly the sense of replacement specified by the model. However, replacement arguments gain their power from the idea of reduction: that is, that the manifest variates can be summarized or replaced by a smaller set of variates. What then is the point of replacing n variates with an infinity? Why would such a replacement criterion have been adopted in the first place? One possibility is that the factor analysis model, just one many possible mathematical phrasings of Spearman's ideas, inadvertently turned out to be a weak phrasing. It must be remembered that a unique replacement was an important feature to Spearman himself, he originally believing that the model *did* provide a unique replacement, and calling Garnett's (erroneous) proof of uniqueness a "momentous theorem" (Spearman, 1927, p. vii). A second possibility is that factor analysis was simply a great trade-off of dimensionality reduction for indeterminacy. It must be remembered that in the factor analysis sense of unidimensionality it is (Σ - Ψ^2), and not Σ (the data), that is of rank one. The

dimensionality of the data is reduced by subtracting Ψ^2 from Σ. Howev the cost of reducing dimensionality in this manner might be viewed as criterion of identity for the replacement variate, that is, "X that is the late common factor to $\mathbf{Y}$", that has a greater latitude with regard to the rando variates it admits than, for example, the criterion found in PCA. In princip component analysis one has large dimensionality and no indeterminacy, a in factor analysis one has minimum dimensionality and indeterminacy.

Conclusion

To paraphrase Hacker (1987), analogy is the go-kart of creativity. T same might said about figurative language in general. However, in scienti and mathematical contexts figurative language has a tendency to get aw from its users if not handled with care. In practice, it is easy to over-exte the reach of a metaphor or analogy. The issue of indeterminacy is a case point: Investigators employing factor analysis routinely speak "underlying", "unobservable", "unmeasurable", "factors responsible for.. And while, in any given substantive problem, it may be the case that the does exist such entities, the factor analysis model does not address this iss Instead, it provides a criterion for the technical concept of *latent comm factor*. When the model describes $\mathbf{Y}$, this criterion admits an infinity random variates that is each *criterially* a "latent common factor to $\mathbf{Y}$". This indeterminacy. And this is clear when metaphor is distinguished from math

References

Anderson, T. & Rubin, M. (1956). Statistical inference in factor analysis. *Proceedings of t Third Berkeley Symposium on Mathematical Statistics and Probability*, *5*, 11-150.

Baker, G. & Hacker, P. (1980). *Wittgenstein: Meaning and understanding*. Chicag University of Chicago Press.

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

Camp, B. (1932). The converse of Spearman's two-factor theorem. *Biometrika*, *24*, 260-26

Garnett, J. (1919). On certain independent factors in mental measurements. *Proceedings the Royal Statistical Society*, *96*, 91-111.

Garrison, J. (1986). Some principles of postpositivistic philosophy of science. *Educatio Researcher*, *15*, 12-18.

Guttman, L. (1955). The determinacy of factor score matrices with implications for five ot basic problems of common-factor theory. *The British Journal of Statistical Psycholo 8* (Part II), 65-81.

Guttman, L. (1977). What is not what in statistics. *The Statistician*, *26*, 81-107.

Hacker, P. (1987). Languages, minds and brains. In C. Blakemore & S. Greenfield (Ed *Mindwaves*. Oxford: Basil Blackwell Ltd.

Hacker, P. (1988). Language, rules and pseudo-rules. *Language and Communication, 8*, 159-172.
Holland, P. (1990). On the sampling theory foundations of item response theory models. *Psychometrika, 55*, 577-601.
Jensen, A. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review, 39*, 1-123.
McDonald, R. (1974). The measurement of factor indeterminacy, *Psychometrika, 39*, 203-222.
McDonald, R. (1979). The simultaneous estimation of factor loadings and scores. *The British Journal of Mathematical and Statistical Psychology, 32*, 212-228.
McDonald, R. & Burr, E. (1967). A comparison of four methods of construction factor scores. *Psychometrika, 32*, 381-401.
Mulaik, S. (1976). Comments on "The measurement of factorial indeterminacy". *Psychometrika*, 41(2).
Mulaik, S. (1990). Blurring the distinctions between component analysis and common factor analysis. *Multivariate Behavioral Research, 25(1)*, 53-59.
Mulaik, S. (1993). The critique of pure statistics: Artifact and objectivity in multivariate statistics. In B. Thompson (Ed.), *Advances in Social Science Methodology, Vol. 3.* Greenwich, CT: JAI Press.
Mulaik, S. (1994). Kant, Wittgenstein, objectivity, and structural equation models. In C. R. Reynolds (Ed.), *Cognitive assessment: A multidisciplinary perspective.* New York: Plenum.
Mulaik, S. & McDonald, R. (1978). The effect of additional variables on factor indeterminacy in models with a single common factor. *Psychometrika, 43*, 177-192.
Piaggio, H. (1931). The general factor in Spearman's theory of intelligence. *Nature, 127*, 56-57.
Schönemann, P. (1990). Facts, fictions and common sense about factors and components. *Multivariate Behavioral Research, 25(1)*, 47-51.
Schönemann, P. & Wang M. (1972). Some new results on factor indeterminacy. *Psychometrika, 37*, 61-91.
Spearman, C. (1927). *Abilities of man.* New York: Macmillan.
Spearman, C. (1933). The uniqueness of *g*. *Journal of Educational Psychology, 24*, 106-108.
Steiger, J. (1979). The relationship between external variables and common factors. *Psychometrika, 44(1)*, 93-97.
Steiger, J. (1990). Some additional thoughts on components, factors, and factor indeterminacy. *Multivariate Behavioral Research, 25(1)*, 41-45.
Steiger, J. & Schönemann, P. (1978). A history of factor indeterminacy. In S. Shye (Ed.), *Theory construction and data analysis in the social sciences.* San Francisco: Jossey Bass.
Ter Hark, M. (1990). *Beyond the inner and the outer: Wittgenstein's philosophy of psychology.* London: Kluwer.
van der Leeden, R. (1990). *Reduced rank regression with structured residuals.* Leiden: DSWO Press.
Vittadini, G. (1989). Indeterminacy problems in the LISREL model. *Multivariate Behavioral Research, 24(4)*, 397-414.
Williams, J. (1978). A definition for the common factor analysis model and the elimination of problems of factor score indeterminacy. *Psychometrika, 43*, 293-306.
Wilson, E. (1928). On hierarchical correlation systems. *Proceedings of the National Academy of Science, 14*, 283-291.
Wittgenstein, L. (1953). *Philosophical investigations.* Oxford: Basil Blackwell Ltd.

Accepted March, 1996.

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 (199
and also helps dispel some of the myths connected with the *posterior moment position.*

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

In this commentary, I present a simple numerical demonstration, (a) an aid to those unfamiliar with the intricacies of indeterminacy, and (b) as challenge to some who *are* familiar, and claim indeterminacy has important consequences. This numerical example seems to demonstra rather convincingly, that some of the assertions of the proponents of PMP a incorrect.

A Signal from Space?

Suppose it is *believed* that a signal may be emitted ten times (*and on ten times*) at one hour intervals from point X, starting at 1:00 P.M. Point can never be observed directly. Receivers are constructed to measure t signal at points Y_1, Y_2, Y_3, Y_4. However, the signal is "jammed" by a noi countersignal at each receiving point. An additional signal, at point W, received directly without any noise degradation. Fortunately, it is know from intelligence sources that (a) all signal and noise distributions have ze means, (b) the signal and noise components are additive, and (c) they a