Michell, J. (2000). Normal science, pathological science and psychometrics. *Theory & Psychology*, *10*, 639–667.
Michell, J. (2004). Item response models, pathological science and the shape of error: Reply to Borsboom and Mellenbergh. *Theory & Psychology*, *14*, 121–129.
Michell, J. (2008). Conjoint measurement and the Rasch paradox: A response to Kyngdon. *Theory & Psychology*, *18*, 119–124.
Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.
Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422.
Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability and font difficulty on the eye movements of young and older readers. *Psychology and Aging*, *21*, 448–465.
Shankweiler, D., & Crain, S.(1986). Language mechanisms and reading disorder: A modular approach. *Cognition*, *14*, 139–168.
Stenner, A. J., Burdick, H., Sanford, E. E., & Burdick, D. S. (2006). How accurate are Lexile text measures? *Journal of Applied Measurement*, *7*, 307–322.
Stenner, A. J., Smith, M., & Burdick, D. S. (1983). Toward a theory of construct definition. *Journal of Education Measurement*, *20*, 305–315.
von Eye, A. (2005). Review of Cliff and Keats, *Ordinal measurement in the behavioral sciences. Applied Psychological Measurement*, *29*, 401–403.

Manifest and Latent Variates

Michael D. Maraun and Peter F. Halpin
Simon Fraser University

In focusing their attentions on latent variable theory, the semantics of constructs, and measurement theory, respectively, Professors Borsboom, Markus, and Michell could not have chosen topics closer to the methodological nerve center of the behavioral and social sciences. We regret having so little space to address the fascinating contents of the current contributions. On the other hand, the appearance of Professor Borsboom's article on latent variable theory is fortuitous, for this topic is particularly dear to our hearts, the first author having just completed 8 years of work on a manuscript, *Myths and Confusions: Psychometrics and the Latent Variable Model*, whose focus is this very issue.

Our recent work on Meehl's taxometric procedures necessitated that we sketch out the full nature of the dependency of the practice of latent variable modeling on the model of detection theory as it is implemented within the natural

Correspondence should be addressed to Michael D. Maraun, Department of Psychology, Simon Fraser University, 888 University Drive, Burnaby, V5A 1S6, Canada. E-mail: maraun@sfu.ca

sciences (see, especially, Maraun, Halpin, Slaney, Gabriel, & Tkatchouk, 2007; Maraun & Halpin, 2007). On this model, latent variable models are portrayed as tools that can be used to detect unobservables (latent structures, causal mechanisms, abstractive properties). The researcher's task is characterized as being to assemble a set of *indicators* on the basis of which, under certain conditions, the latent variable model can *tap into* (a metaphor drawn from the domain of resource detection and extraction) the *latent realm* and access *underlying* unobservables. As argued extensively in *Myths and Confusions*, this account is a mythology, called therein the *Central Account*. Quarks, neutrinos, metallic objects, alpha particles, etc., are existing constituents of natural reality, and for such entities, detectors can and have been built. In contrast, there has never been specified any class of constituents of natural reality for which the latent variable model could play the role of detector. But then, latent variable models are not *about* unobservables, or any other class of constituents of natural reality, but, rather, random variates and their properties.

As with Rod McDonald's (1974) earlier attempt to capture the essential difference between the two types of variates that are featured in the equations of latent variable models, Borsboom's attempt, which rests on the notion of epistemic accessibility, *presupposes* the mythology of the Central Account and thus must answer to it. But that which sails on the wind produced by a mythology is destined to run aground on the shores of incoherence and confusion, and Borsboom's account is no exception. In particular, in order to service a mythology, Borsboom is forced to deform the core components of scientific investigation: concepts, their referents, variables, models, inference, definition, etc. Neither the concept *latent variable* nor particular latent variables are "hypothetical structures." *Sex* and *age* are not variables but rather concepts. It makes no sense to "conceptualize extraversion as a latent variable" (Borsboom, 2007, p. 8) because the concept *extraversion* is a concept, not a variable. Concepts are constituents of language whose correct employments are fixed by linguistic rules. The concept *extraversion* denotes certain, particular psychological phenomena. Although it is a hallmark of science to clarify the correct employments of concepts that denote and organize phenomena to be studied (think, e.g., of Mach on Newton's conception of mass, Einstein on simultaneity), the Central Account convinces the latent variable modeler to bypass this work. It is as if by writing down the symbol θ and incanting *extraversion*, a link can be forged between θ and the linguistic superstructure that invests the concept *extraversion* with meaning.

The referents of certain concepts are constituents of natural reality that are extended in space, and these constituents, with respect some particular observational setup, can be coherently placed on the dimension of observable to unobservable. A variable, on the other hand, is simply a rule/map/function (see Rozeboom, 1988; Maraun, 2007) and, hence, cannot coherently be placed on the dimension of observable to unobservable. The fact that a researcher can make errors of measurement or errors in inference does not imply that anything is "underlying": This is simply the mythology of the Central Account speaking. Most essentially, it makes no sense to claim, as does Borsboom (2007, p. 9), that "when we treat a variable as 'observed' we mean nothing more than that we assume that the location of a person on that variable can be inferred with certainty from the data." The Central Account here visits upon Borsboom's account the ghostly shapes of underlying, latent realms, replete with their platonic contents. But science does not work this way. Variables are simply functions, not constituents of natural reality. Variables are constructed by scientists to map phenomena onto the real line and, hence, they *produce* the data that the scientist uses to study this phenomenon. It is, thus, incoherent to talk of inferring the location of a person on a variable from the data: To know a person's score on a variable does not require an inference of any sort but simply knowledge of the variable (i.e., the rule that maps a person from some contrast class onto the real line).

The clue to what latent variable models are, and to a workable account of the basis for the traditional manifest/latent variable distinction, lies in a reconsideration of the indeterminacy property of linear factor structures. Of necessity, what we offer here is but a sketch of a case argued elsewhere (Maraun, 2007) in great detail. Latent variable models are not detectors of unobservable latent structures, properties/attributes, causal sources, or anything else. They are of a kind with component models, and, while both possess model-like features, neither is a model in the classical sense of the term (for neither involves rules that establish entity/symbol correspondences). Both component and latent variable models are really *replacement variate generators*, quantitative recipes for the construction of variates, $\boldsymbol{\theta}_k$, $k = 1 \ldots m$, that *replace* a set of variates $\mathbf{X}_j$, $j = 1 \ldots p$, in some particular, optimal fashion. Latent variate generators involve two types of variate, the *input variate* (traditionally, the manifest variable) and the *replacement variate* (traditionally, the latent variable). Note that input and replacement variates are *variates*, hence, simply rules/maps/functions. Once the mythology of the Central Account is held at bay, the actual basis for the classical manifest/latent variable distinction then becomes clear.

The scores that comprise the distributions of the p input variates, $\mathbf{X}_j$, $j = 1 \ldots p$, are produced in accord with *rules of score production*, r_j, $j = 1 \ldots p$, *known by the researcher prior to analysis*. The fact that scores can be produced for each input variate prior to analysis by taking as the argument of r_j each member, p_i, of some set P of objects under study, i.e., $\mathrm{x}_{ji} = \mathrm{r}_j(\mathrm{p}_i)$, $\mathrm{p}_i \in P$, explains why scores on these variates constitute the data to be analyzed. In contrast, there does not exist, prior to analysis, rules by which the scores that comprise the distributions of the replacement variates, $\boldsymbol{\theta}_k$, $k = 1 \ldots m$, can be produced. This is why scores on replacement variates (latent and component variates) are *not* part of the data to be analyzed.

Let the equations of a given replacement variate generator, R, be symbolized as f($\underline{\mathbf{X}}$; $\underline{\boldsymbol{\theta}}$), in which $\underline{\mathbf{X}}$ is a *p*-element vector of input variates and $\underline{\boldsymbol{\theta}}$, an *m*-vector of replacement variates. Let any additional distributional and moment constraints inherent to R be symbolized as D, and let $\underline{\Pi}$ contain the parameters of the generator. A given replacement variate generator can then be symbolized in full as R:[f($\underline{\mathbf{X}}$; $\boldsymbol{\theta}$),D,$\underline{\Pi}$]. In the special case of a unidimensional generator (hereafter, exclusively considered), $\underline{\boldsymbol{\theta}}$ is a single variate $\boldsymbol{\theta}$. The symbol $\boldsymbol{\theta}$ is a placeholder for *any* variate constructed so as to satisfy R:[f($\underline{\mathbf{X}}$; $\boldsymbol{\theta}$),D,$\underline{\Pi}$], when, in fact, $\underline{\mathbf{X}}$ satisfies it. In the event that particular $\underline{\mathbf{X}}$ satisfies the requirements stipulated by R:[f($\underline{\mathbf{X}}$; $\boldsymbol{\theta}$),D,$\underline{\Pi}$], it can be said that R:[f($\underline{\mathbf{X}}$; $\boldsymbol{\theta}$),D,$\underline{\Pi}$] *describes* $\underline{\mathbf{X}}$. Any variate $\boldsymbol{\theta}$ that satisfies R:[f($\underline{\mathbf{X}}$; $\boldsymbol{\theta}$),D,$\underline{\Pi}$] can be called a *replacement variate to* $\underline{X}$ *(under R:[f(*$\underline{X}$*;**r**),D,*$\underline{\Pi}$*])* because it stands in place of, or approximates, $\underline{\mathbf{X}}$, in precisely the sense specified by R:[f($\underline{\mathbf{X}}$; $\boldsymbol{\theta}$),D,$\underline{\Pi}$].

In describing, and distinguishing between, replacement variate generators, the following issues are relevant. *Existence*: For given replacement variate generator, R:[f($\underline{\mathbf{X}}$; $\boldsymbol{\theta}$),D,$\underline{\Pi}$], what conditions must $\underline{\mathbf{X}}$ satisfy in order that there exists at least one variate $\boldsymbol{\theta}$ that replaces it in the sense specified by this generator, i.e., in order that $\underline{\mathbf{X}}$ be replaceable under R:[f($\underline{\mathbf{X}}$; $\boldsymbol{\theta}$),D,$\underline{\Pi}$]? *Cardinality of replacement*: If $\underline{\mathbf{X}}$ is replaceable under R:[f($\underline{\mathbf{X}}$; $\boldsymbol{\theta}$),D,$\underline{\Pi}$] , what is the cardinality of the set *C* of replacement variates $\boldsymbol{\theta}$? *Construction formulas*: Let T represent the totality of requirements imposed on $\boldsymbol{\theta}$ by R:[f($\underline{\mathbf{X}}$; $\boldsymbol{\theta}$),D,$\underline{\Pi}$] when it describes $\underline{\mathbf{X}}$, i.e., the totality of requirements that a variate must satisfy in order for inclusion in *C*. Does there exist a construction formula, $\boldsymbol{\theta}$ =T[R_1:[f($\underline{\mathbf{X}}$;) $\boldsymbol{\theta}$,D,$\underline{\Pi}$]], according to which each of the elements contained in *C* can be produced?

As an example, consider the unidimensional, linear factor (ulf) generator. Let $\underline{\mathbf{X}}$ be a *p*-vector of input variates for which $E(\underline{\mathbf{X}}) = \underline{0}$ and $E\underline{\mathbf{X}}\underline{\mathbf{X}}' = \Sigma$. A continuous variate $\boldsymbol{\theta}$ is sought that possesses the following properties: (ri) $E(\boldsymbol{\theta}) = 0$; (rii) $E\boldsymbol{\theta}^2 = 1$; (riii) the vector of residuals, $\underline{\mathbf{l}}$, of the linear regression of $\underline{\mathbf{X}}$ on $\boldsymbol{\theta}$ has a covariance matrix that is diagonal and positive definite. *Existence*: If $\Sigma = \underline{\Lambda}\underline{\Lambda}' + \Psi$, Ψ diagonal and positive definite, then there exists a $\boldsymbol{\theta}$ that satisfies (ri), (rii), and (riii) (Wilson, 1928; Guttman, 1955), i.e., $\underline{\mathbf{X}}$ is ulf-replaceable. *Cardinality of replacement*: If $\underline{\mathbf{X}}$ is ulf-replaceable, then $\mathrm{Card}(C) = \infty$ (Wilson, 1928; Piaggio, 1931). That is, if $\underline{\mathbf{X}}$ is ulf-replaceable, then its replacement is not unique, there being constructible an infinity of random variates each of which satisfies (ri)–(riii). *Construction formula*: As distinct from component replacement variates, $\boldsymbol{\theta}$ cannot be constructed as a linear function of $\underline{\mathbf{X}}$ (e.g., McDonald, 1977; Maraun, 2007). If $\underline{\mathbf{X}}$ is ulf-replaceable, the construction formula for ulf replacement variates $\boldsymbol{\theta}$ (the infinity of elements of set *C*) is $\theta = \underline{\Lambda}'\Sigma^{-1}\underline{\mathbf{X}} + w^{1/2}\mathbf{s}$, in which $w = (1 - \underline{\Lambda}'\Sigma^{-1}\underline{\Lambda})$ and $\mathbf{s}$ is an arbitrary random variate for which $C(\mathbf{s},\underline{\mathbf{X}}) = \underline{0}$, $E(\mathbf{s}) = 0$, and $V(\mathbf{s}) = 1$ (Piaggio, 1931; Kestelman, 1952; Guttman, 1955).

Mythology-free accounts of a number of the most frequently employed replacement variate generators, including the principal component, LISREL, and two-parameter item response generator, have been worked out in Maraun (2007).

REFERENCES

Borsboom, D. (2008). Latent variable theory. *Measurement: Interdisciplinary Research and Perspectives*, *6*, 25–53.

Guttman, L. (1955). The determinacy of factor score matrices with implications for five other basic problems of common-factor theory. *The British Journal of Statistical Psychology*, *8*(2), 65–81.

Kestelman, H. (1952). The fundamental equation of factor analysis. *British Journal of Psychology*, *5*, 1–6.

Maraun, M. (2007). *Myths and confusions: Psychometrics and the latent variable model*. Retrieved from http://www.sfu.ca/~maraun/Mikes%20page-%20Myths%20and%20Confusions.html

Maraun, M., & Halpin, P. (2007). *MAXEIG: An analytical treatment*. Retrieved from http://www.sfu.ca/~maraun/Mikes%20page-%20Ongoing%20Research.html

Maraun, M., Halpin, P., Slaney, K., Gabriel, S., & Tkatchouk, M. (2007). *What the researcher should know about Meehl's taxometric tools of detection*. Retrieved from http://www.sfu.ca/~maraun/Mikes%20page-%20Ongoing%20Research.html

McDonald, R. P. (1975). Descriptive axioms for common factor theory, image theory and component theory. *Psychometrika*, *40*(2), 137–152.

McDonald, R. P. (1974). The measurement of factor indeterminacy. *Psychometrika*, *39*, 203–222.

Piaggio, H. (1931). The general factor in Spearman's theory of intelligence. *Nature*, *127*, 56–57.

Rozeboom, W. (1988). Factor indeterminacy: The saga continues. *British Journal of Mathematical and Statistical Psychology*, *41*, 209–226.

Wilson, E. B. (1928). On hierarchical correlation systems. *Proceedings of the National Academy of Science*, *14*, 283–291.

Lost in Translation? Meaning and Decision Making in Actual and Possible Worlds

André A. Rupp
University of Maryland

From different angles, Borsboom, Markus, and Michell present a careful analysis of the way that specialists reason with empirical data about latent characteristics of individuals. They jointly argue for a more precise and thoughtful use of the key

Correspondence should be addressed to André A. Rupp, Department of Measurement, Statistics, and Evaluation (EDMS), 1230 Benjamin Building, University of Maryland, College Park, MD 20742. E-mail: ruppandr@umd.edu