A Descriptivist Approach to Trait Conceptualization and Inference

Katherine G. Jonas

University of Iowa

Kristian E. Markon

University of Iowa

Abstract

In their recent article, “How Functionalist and Process Approaches to Behavior Can Explain Trait Covariation,” Wood, Gardner, and Harms (2015) underscore the need for more process-based understandings of individual differences. At the same time, the article illustrates a common error in the use and interpretation of latent variable models: namely, the misuse of models to arbitrate issues of causation and the nature of latent variables. Here, we explain how latent variables can be understood simply as parsimonious summaries of data, and how statistical inference can be based on choosing those summaries that minimize information required to represent the data using the model. Although Wood, Gardner, and Harms acknowledge this perspective, they underestimate its significance, including its importance to modeling and the conceptualization of psychological measurement. We believe this perspective has important implications for understanding individual differences in a number of domains, including current debates surrounding the role of formative versus reflective latent variables.

Keywords: latent variable, description length, trait theory, model inference

Introduction

We read with interest the work of Wood, Gardner, and Harms (2015, "How Functionalist and Process Approaches to Behavior Can Explain Trait Covariation"). Direct assessment of abilities, expectancies, and valuations deepens our understanding of human behavior, and strengthens the framework of process-oriented accounts of individual differences. In this regard, we welcome the functionalist paradigm outlined by Wood, Gardner and Harms (2015), and believe it represents an important direction forward.

At the same time, their treatment of one theme—the interpretation of structural models—highlights a number of misconceptions regarding the use of latent variable models in studying individual differences. Here, we discuss two particular misunderstandings: one is the erroneous assumption that latent variables cannot represent summaries of behavior; the second concerns how to define parsimony. We demonstrate that the interpretation of traits as summaries does not imply the models generally suggested in the literature, including those used by Wood, Gardner, and Harms (pp. 86, 104). Second, although it may not "*always* [be] obvious how to judge" parsimony (p. 104, emphasis added), there is a large domain of literature explaining how to empirically judge the parsimony of statistical models, at least. We briefly outline a descriptivist account of latent variables and statistical inference as it applies to psychological measurement and modeling. We borrow the term "descriptivist" from Myung, Navarro, and Pitt (2006), who applied it more generally to all statistical models and inference. In the descriptivist paradigm, latent variables are interpreted simply as summaries of data, and the parsimony of those summaries is the metric for model comparison. This interpretation of latent variables dismantles the statistical distinction between Wood, Gardner, and Harms' proposal and existing approaches to personality, providing a framework for integrating the two.

The Characteristics of Latent Variables

[Figure 1 about here]

Reflective and formative latent variable models. Wood, Gardner and Harms distinguish two approaches to the interpretation of latent variable models of personality, which they term the structuralist and functionalist approaches. The authors refer to the structuralist approach as "appealing to broad factors identified by investigations of trait structure" (p. 84), and associate it with the use of reflective latent variable models. An example of a reflective latent variable model is shown in Figure 1a (see also Figure 1a of Wood, Gardner, and Harms), where the item responses x_i are a function of the latent trait η, the item loading λ_i, and a random error term for each item ε_i. In diagrams of these models, the paths flow from the latent variables to the observed indicators.

Wood, Gardner and Harms contrast the structuralist perspective with their proposal, the functionalist approach, which focuses on specific functions of behavioral traits, such as a person's abilities, expectancies, and valuations. In their proposal, functionalist items pertaining to a given behavioral trait are used to construct formative latent variable models, an example of which is shown in Figure 1b. In diagrams of these models, the paths flow from the observed indicators to the latent variables. In Figure 1b, γ_i is the loading of the i^{th} item on the latent trait, the y_is are endogenous variables, ζ is a disturbance term, and the other terms are defined as in the reflective latent variable model. In the framework proposed by Wood, Gardner, and Harms, functionalist items would make up the x_is in Figure 1b.

Equating structuralist approaches with reflective models and functionalist approaches with formative models, as Wood, Gardner, and Harms do, is unwarranted. Although structuralist accounts of personality have historically employed reflective latent variable models, and have often interpreted latent variables as unitary causative factors (Cattell, 1950, and others discussed on pages 85-86 of Wood, Gardner, and Harms), the association between the method and its interpretation is an artifact of historical trends within this particular area of study. As Wood, Gardner, and Harms acknowledge (e.g., pgs 86, 100), structuralist questions are orthogonal to the methods typically used to answer them.

We agree, and further suggest that some of the features of the functionalist theory Wood, Gardner, and Harms propose are more accurately represented by traditional reflective latent variable models. In particular we argue that, contrary to what is assumed in the literature, the descriptive, summary interpretation of traits (ascribed to functionalism by Wood Gardner and Harms) is in fact better represented by reflective latent variable models than formative latent variable models. To delineate this perspective, we outline what we term the descriptivist approach, in which latent variables are understood as parsimonious summaries of data—no more and no less. The descriptivist approach has been discussed previously in the psychological literature (e.g., Myung, Navarro, & Pitt, 2006), but is not as widely known as it should be, especially in the personality and measurement literature.

Must reflective latent variables be real, causative, or unitary? Wood, Gardner and Harms imply that a reflective latent variable model is inconsistent with a descriptive, non-causal interpretation of traits. This is likely a function of the the widespread, but erroneous assumption that the very form of the reflective latent variable model requires that a latent variable be interpreted as a physical etiologic agent of the observed data (Borsboom, Mellenbergh, & van Heerden, 2003). In practice, however, the use of latent variable models is often incompatible with such an ontology. Latent variables can be used descriptively, as shorthand method of representing correlations among observed behaviors. Latent variable models of infant attachment

(Kroonenberg & van Dam, 1997), for example, describe a theoretical construct that exists only in the space between two individuals. The latent attachment variable is not conceptualized as real, in the sense of having some one-to-one correspondence with a physical etiologic agent.

Reflective latent variables need not be unitary, either. Behavioral genetic research, for example, invokes latent variables as a way to describe genetic liabilities to psychopathology (Kendler KS, Prescott CA, Myers J, & Neale MC, 2003). A latent variable in a behavioral genetic model is understood to represent an agglomeration of etiologic effects (e.g., polymorphism effects), rather than a unitary, homogenous entity.

Reflective latent variables need not be causal with respect to their indicators, either. A model of correlated intelligences (van der Maas et al., 2006) demonstrates that a single reflective latent variable can describe data generated from multiple factors. Van der Maas and colleagues proposed that the observed correlations between cognitive abilities, such as verbal and spatial abilities, arise from a mutualism model. The mutualism model, which was first used as a model of ecosystem and societal development, specifies that cognitive abilities are initially uncorrelated, but grow cooperatively—that is, abilities in one area tend to aid development of other abilities. The resulting data is best described by a single reflective latent variable, consistent with a general intelligence factor. However, the single factor is not the causative agent, even though it describes the data better than any alternative model.

Similarly, Sampson (1968) collected a dataset consisting of monks' report of their positive relationships with other monks, and interpreted the resulting network as indicative of four groups. Hoff, Raftery, and Handcock (2002) replicated Sampson's four groups by applying a latent space model—which can be shown to be a reflective latent variable model—to the data. Sampson defined the groups as "outcasts," "young turks," "waverers," and "the loyal opposition," based on his observations of the monks' political leanings, and indeed, political allegiances may be a causative factor. But the causative agent could just as well be shared mindsets or personality traits, rooming arrangements, age, or a combination of these factors, none of which are necessarily isomorphic with the factors used to define the latent space. By analogy, high school school cliques are not causal with respect to the relationships among their members—two "popular" girls are not necessarily friends *because* of their latent group membership, just as a lack of positive relationships is not caused by being a member of a "loner" group. The relationships, or lack thereof, are more probably a function of shared interests, age, physicality, etc. The latent groups are just a convenient way of summarizing patterns of observed relationships.

Reflective latent variable models are agnostic with regard to the nature of the etiological process: this is the heart of the descriptivist paradigm. Reflective latent variable models can describe data generated from any number of etiologic processes; consequently, the form of the latent variable model cannot arbitrate questions of causality. Doing so would require manipulating the latent variable, and it is arguably just as difficult to manipulate formative variables—for example, the prototypical formative variable, socioeconomic status—as reflective variables.

The descriptivist approach allows for the adoption of latent variables without committing to any assumptions about their nature. In fact, the descriptivist paradigm questions the meaningfulness of doing so. Latent variables are simply model parameters that represent subsets of the data. Consequently, as we describe below, statistical inference about latent variables becomes a search for the models that subset the data most parsimoniously. Contrary to the suggestion of Wood, Gardner, and Harms (2015, p. 104), this search can be carried out in an empirical, mathematically rigorous way (Cover & Thomas, 1991; Rissanen, 1996, 2001).

The Descriptivist Approach

Modeling as Parsimonious Data Description, via Information Theory

The fundamental concept underlying descriptivism is that statistical inference can be seen as a search for models that provide the shortest representation, or description, of data possible. In an information-theoretic sense, a model provides a code for the data, just as a raw string of values is a code. An ideal model is more parsimonious than the data string it encodes, and, although not actually attainable in practice, perfectly encodes the data. When we say we have learned something about the data, we mean we have found a way to represent it using less information.

The link between model inference and data description length is based on the observation that codes for data should be assigned such that code lengths are inversely proportional to the probabilities of the data they encode—that is, the most efficient way to assign codes is to give the shortest code to the most common value, and vice versa. In fact, the shortest code for a data value $\mathbf{x}$, using a model M with parameter value $\boldsymbol{\theta}$, is given by (e.g., Cover & Thomas, 1991):

$$ln[p(\mathbf{x}|M(\mathbf\theta))]$$

meaning that finding the estimate of $\boldsymbol{\theta}$ that maximizes the likelihood of the data minimizes the length of the code (where the units of length are given by the logarithm base, nats if base e, or bits if base 2).

The model itself also has a codelength, which is a function of the sum of maximum likelihoods over all possible datasets (Rissanen, 2001):

$$ln[\displaystyle{\Sigma_y}p*(\mathbf{y}|M)]$$

where $p^*(\boldsymbol{y}|M)$ is the maximum likelihood of the data ***y*** given model *M*, and the sum is taken over all possible datasets. The codelength of a model is therefore proportional to its tendency to produce large likelihoods in general. This reflects an intuitive idea: more complex models will fit any dataset well, not just the observed data. Complex models also require more information to code them. In sum, the model codelength will be larger for models with more parameters, and for models that tend to produce large maximum likelihoods for any data, not just the observed data (e.g., Grünwald, 2007).

Altogether, the total information needed to represent a dataset ***x*** with model M is the information required to code the data given a model (Equation 1), plus the information required to code the model (Equation 2):

$$-\ln[p^*(\mathbf{x}|M)]+\ln[\displaystyle{\Sigma_y}p^*(\mathbf{y}|M)]$$

When we use total information to choose a model, we make a choice based on the models' ability to reproduce the observed data, adjusted for their tendency to reproduce any data well. Models that require less information to do so are more parsimonious, and therefore preferable.

Latent Variables as Representations of Data Subsets: An Example

To illustrate this approach to model comparison and inference, we model the occurrence of positive daily events from the National Survey of Midlife in the United States II (MIDUS II), a large-scale, population-representative longitudinal study of adults in the US. As part of this study, individuals were interviewed daily about positive events for eight days; we focus our discussion on the binary variable of whether or not a positive event was reported each day (743 individuals provided complete data; 56% female; mean age = 58.78, sd = 12.45; for further details, see Ryff & Almeida, 2010; Almeida, Wethington, & Kessler, 2002).

Probability as representation: The binomial model. One way to represent the data is as a string of 5944 individual values, each representing whether a positive event was reported on a given day by a given individual. This representation—the representation one might observe in a database—is very precise, but also relatively lengthy and complex, as it comprises an implicit model uniquely describing each observation—that is, a model with 5944 parameters.

Rather than represent each event individually, one can represent the event variable with a simple model, the binomial model with a single probability parameter. This model represents the 5944 individual values by reencoding them using a single parameter, the probability of an individual experiencing a positive event on a given day, π (the maximum likelihood estimate of π is 0.71). This is a relatively simple model—according to Equation 2, 4.57 nats are used to encode the model (Drmota & Szpankowski, 2004; Rissanen, 1996)—which is intuitively sensible, as the model involves a single parameter that directly provides the expected distribution of the data. Using this model, 3586.54 nats of information are used to encode the data, meaning that the total amount of information needed to encode the data plus the model is 3591.11 nats.

Latent variables as representations of subsets of observations: The random-effects binomial model. In the case of the traditional binomial model, a single parameter is used to represent the entire set of 5944 observations. Another possibility is to subset the data by individual, and use a separate probability parameter to represent each individual's set of responses (that is, to model π_i instead of π, a random-effects binomial). Each π_i represents the probability of individual i reporting a positive event on any given day. Because there are only 9 possible values of π_i (corresponding to 0, 1, 2, … 8 events over the observational period), this model reencodes the 5944 original observations or 743 vectors using 8 different parameter values (because the π_i sum to one).

The model assumes that the number of positive events reported over 8 days is an imperfect estimate of π_i. In other words, as a random-effects binomial model, it is also a reflective latent variable model.[1] It is important to note, however, at no point have we invoked the idea that the latent variable represented by π_i is causal, or that it has any ontological status beyond it being a description of the ith subset of observations. The data could just as well be subset by days, and π_i of the ith day would function similarly in the model as the π_i of the ith person. There may be a real mechanism generating the data, but the model is agnostic to that; the parameters π_i do not differ in their ontological status from the parameter π. Stated differently, the parameters π_i serve the same role as formative variables, summarizing responses from the ith individual, or the ith day. Contrary to what has been assumed in the measurement literature (e.g., Bollen & Bauldry, 2011; Edwards & Bagozzi, 2000; Wang, Engelhard, & Lu, 2014; West & Grimm, 2014), this *reflective* latent variable *summarizes* the set of 8 observations for each individual.

As would be expected, this random-effects binomial model is more complex than a simple fixed-effects binomial model, requiring 275.06 nats to encode the model. However, only 3269.93 nats are needed to encode the data using this model, less information than is required using the fixed-effects binomial. Together, the information required to encode the data plus the model is 3544.99 nats, meaning that the even though the random-effects binomial is more complex than its fixed-effect counterpart, it affords a shorter description of the data, and summarizes it more parsimoniously. From a classical inferential perspective, this is equivalent to concluding that the random-effects binomial model—that is, the reflective latent variable model—fits the data better than the fixed-effect model.

---

[1] The random-effects binomial can be shown to be a form of parallel-indicators item response theory (IRT) model by setting $\pi_i = [1+exp(-\theta_i)]^{-1}$ (e.g., Curran, 2003).

Formative variables as deterministic constructions: Evaluating hypotheses about construct identity. What would it mean to create a formative variable from the daily events, as Wood, Gardner, and Harms do for personality measures (e.g, Figures 1 and 2; pp. 86, 104)? One could consider the sum of positive events over the eight days, and this would be perfectly correlated with the maximum likelihood estimate of π_i under the reflective, random-effects binomial model. So, in this case, is a formative variable approach equivalent to a reflective latent variable model?

A critical difference between formative and reflective variables is that the latter specify the probability of the data, whereas the former do not. That is, a reflective model allows one to calculate the likelihood that person reports x positive events, given their standing on the latent trait, π_i. For example, according to the probability mass function of the binomial distribution, the likelihood that an individual with π_i=.5 reports four positive events over the eight days is .27. In contrast, a formative model does not allow for this calculation, because formative variables are linear combinations of observed variables. If an individual reports four positive events over eight days, π_i is 4, but there is no way of backtracking to quantify the likelihood of the observed data, because π_i is completely determined by x.

As a result, it is not possible to evaluate equations 1 or 3 for formative variables, making it impossible to quantify model parsimony, and consequently, to make empirical inferences between models. Reflective variables can be empirically compared on the basis of their parsimony (Equations 1 and 3), and formative variables cannot. There is a probabilistic model implied by the formative paradigm, but it is invisible, lying to the left of the the x_is in Figure 1b. Returning to the daily events example, if a formative variable were constructed from the observed daily responses (the x_is), the distribution of the formative variable η would be determined by the observed responses (the terms to the right of the formative variable would not exist). The actual model describing the distribution of the observed data lies, implicitly, to the left of the x_is, and could be any number of models.

Technically, purely formative models do not exist, as they are underidentified, a point that has been made repeatedly in the literature (the latent disturbance term ζ in Figure 1b is identified only when variables $y_{1\text{-}3}$ are included in the model; Howell, Brievik, & Wilcox, 2007; Lee, Cadogan, & Chamberlain, 2013). A purely causative formative structure (e.g., one having disturbances without also having a reflective latent variable) is not identifiable, and therefore has no estimable parameters; a composite formative structure (without disturbances) has no measurement parameters to estimate or make inferences about.

In the same vein, principal components are are often thought of as formative constructs in which the item weights (γ_i in Figure 1b) are determined by the items' observed correlations with each other (cf. Footnote 1 of Wood, Gardner, and Harms, 2015). However, even principal components become reflective when they are converted to summaries. Principal components only represent reexpressions of observed data when one extracts as many components as variables—that is, when no components are dropped. When components are dropped, as is standard practice, the dropped components are assumed to be uninterpretable, random noise. In such cases, the components analysis becomes equivalent to an isotropic error factor analysis model (i.e., a FA model where error variances are constrained to be equal; Tipping & Bishop, 1999), and the summarizing components become equivalent to latent variables.

Measurement Error as the Information Cost of Using Descriptive Summaries

A key distinction between reflective and formative variables is that reflective variables, because they include a random error term for each indicator (ε_i in Figure 1a), acknowledge and estimate measurement error. Formative variables, in their pure form, do not (Diamantopoulos, 2006). From an information-theoretic perspective, measurement error can be thought of as the information cost of using a latent variable to represent the data. If a reflective latent variable is a summary, it should entail some loss of information relative to the original data.[2] The benefit of using a latent variable is that it is transferable across observational occasions, and in this sense is more parsimonious than not using any model at all, which would require a new model for each dataset. For example, when we use a trait estimate from a test, we lose information contained in the item responses, but the trait estimate is more parsimonious than treating each item response as independent. That is, we gain parsimony and generalizability by using the latent variable, but it comes at the cost of extra information to encode the data. This cost is proportional to the measurement error.

The shortest code for a vector of responses is given by its empirical probability, or maximum likelihood under a saturated model $p^*(\boldsymbol{x})$ (cf. Equation 1). Given that the information required to encode this is -ln[$p^*(\boldsymbol{x})$], and the information required to encode it using the latent variable model is -ln[$p^*(\boldsymbol{x}|M(\boldsymbol{\theta}))$], then the extra information required to code the responses using the model, on average, is:

$$\displaystyle{\Sigma_x} \big(\pi(\mathbf{x})(\ln[p^*(\mathbf{x})]-\ln[p^*(\mathbf{x}|M(\mathbf{\theta}))])\big)$$

where $\pi(\boldsymbol{x})$ specifies the distribution of the data. Returning to the random-effects model of positive events, if an individual's actual probability of reporting a positive event, $p^*(\boldsymbol{x})$, is 0.625, then on average one would have to use an extra 0.55 nats of information to describe their

[2] While “cost” and “loss of information” have negative connotations, some loss of fidelity is required in order to say something meaningful about the data, just as a map communicates contour by sacrificing information about color or vegetation.

responses over the course of 8 days, relative to simply reporting the proportion of days in which they reported a positive event during a given observational period. In other words, 0.55 nats is the informational cost of having a transferable summary in the form of the latent variable π_i.

The same balancing of costs and benefits extends to the multivariate case. When data is normally distributed, the information cost of using latent variables is a function of measurement error (in the form of the item residual variances; Appendix A). As residual variances approach zero, the information cost of using the latent variables is minimized; as the residual variances increase, the information cost increases. Note that we have not assumed that the population correlation matrix actually reflects real latent variables. We only assume that one wishes to use latent variables to summarize the observed data. In this context, residual variances become indices of the information loss that is a necessary byproduct of using latent variables as summary descriptions of behavior. We generally accept this cost, as it is more parsimonious in the long run—that is, when we wish to assume stability in individual behavior and generalize our findings beyond one dataset.

Implications of the Descriptivist Paradigms: A Structuralist Interpretation of Functionalist Measures

The paradigm outlined here provides an alternative interpretation of the framework suggested by Wood, Gardner, and Harms (2015). For example, although the authors acknowledge the summary interpretation of latent variables (e.g., p. 86, 100), they suggest that structural models involving latent variables are somehow at odds with this interpretation—that there is a gap between the theoretical interpretation and modeling framework (pp. 86, 104), and that “the causal arrows should point from the specific traits toward the structural factor” (p. 86). As we have shown, there is no gap. In fact, if one conceptualizes of latent variables as summaries, the most appropriate model is a reflective one, in which the arrows point from the structural factor to the specific traits. The descriptivist paradigm suggests that structuralist and functionalist approaches are not at odds with each other, much less mutually exclusive. However, because the use of formative models prevents model comparison, both structuralist and functionalist constructs must be modeled reflectively in order to integrate the two accounts.

If neither formative nor reflective models can inform questions of causality, what is the role of causal explanation in functionalism? Wood, Garner and Harms hypothesize that functionalist traits explain structuralist traits. Presumably, however, some of the environmental and neurobiological factors influencing responses to the items “like to attract attention” and “like to talk in social situations” (both functionalist items) are the same as those influencing Big Five indicators of sociability. So have the functionalist constructs explained the Big Five, or are they explained by the same etiologic explanatory networks undergirding the Big Five? Without a joint analysis of both structuralist and functionalist traits, it is impossible to discern how the functionalist and Big Five indicators differ.

It is also important to consider Wood, Gardner, and Harms’ (2015) criticism of aggregated constructs, in light of the descriptivist approach outlined here. For example, on page 102 the authors note that there is value in examining disaggregated constructs and indicators, observing that “it should be a rare circumstance when all traits to identify a structural factor are equally related with any particular outcome of interest.” We certainly agree, while noting that there is value in using summary constructs as well, for theoretical as well as statistical reasons. When we say a tall person has hit their head on a doorframe, we don’t say say they hit their head because of the dynamic system involving all of the molecular and energetic interactions between the different locations on the skull, scalp, and body on the one hand, and the environment on the other—we explain it in terms of height. Sometimes there is value in aggregation, sometimes there is value in disaggregation. The descriptivist asks “how is it most parsimonious to do so?” and “how much information is lost?” Again, without a direct comparison of structural models that comprise functionalist traits, it is not possible to determine which provides the most useful shorthand.

Finally, it is important to comment on the psychometric implications of functionalism, given the discussion above, about measurement error in formative and reflective accounts. As we have noted, the formative framework used by Wood, Gardner and Harms implies no measurement error. Informatically speaking, this suggests that none of the measurements are generalizable. Although pure state variables are of interest in a great number of areas of psychology, in many areas some degree of stability is assumed, and the very purpose of obtaining a measurement of an individual is to generalize that measurement to other occasions. An achievement test, for example, is given precisely because it is believed to measure an attribute that persists beyond a particular time and exam. The utility of functionalist traits is hamstrung by their association with formative latent variable models. In order to achieve generalizability, one must have some sort of summarizing parameter that is modeled and estimated for an individual, implying measurement error, and therefore, a reflective latent variable.

Conclusions

Latent variable models are agnostic with regards to the reality, unity, or causative nature of the latent constructs they measure. We agree with Wood, Gardner, and Harms that “the Big Five structural factors” might provide “useful summaries” of the functionalist indicators (p. 99). According to the descriptivist approach, however, the appropriate method for evaluating this hypothesis is a direct comparison—via reflective latent variables models. In the current context, the descriptivist position highlights how functionalist constructs can be approached from, and integrated with, a structural perspective. More broadly, adopting a descriptivist approach to modeling and inference provides the theoretical foundation for resolving many similar debates in the literature.

References

Almeida, D. M., Wethington, E., & Kessler, R. C. (2002). The daily inventory of stressful events: An interview-based approach for measuring daily stressors. *Assessment, 9*, 41-55.

Bollen, K. A., & Bauldry, S. (2011). Three Cs in measurement models: Causal indicators, composite indicators, and covariates. *Psychological Methods, 16*(3), 265-284. doi:10.1037/a0024448

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*(2), 203–219.

Cattell, R. B. (1950). Describing personality through traits. In R. Cattell (Ed.), *Personality: A systematic theoretical and factual study* (1st ed., pp. 21–44). New York, NY: McGraw-Hill.

Cover, T. M., & Thomas, J. A. (1991). *Elements of information theory*. New York: Wiley.

Curran, P. J. (2003). Have multilevel models been structural equation models all along? *Multivariate Behavioral Research*, *38*, 529-569.

Diamantopoulos, A. (2006). The error term in formative measurement models: interpretation and modeling implications. *Journal of Modelling in Management*, *1*(1), 7–17. doi:10.1108/17465660610667775

Drmota, M., & Szpankowski, W. (2004). Precise minimax redundancy and regret. *IEEE Transactions on Information Theory, 50*, 2686-2707.

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods, 5*(2), 155-174. doi:10.1037/1082-989X.5.2.155

Grünwald, P. D. (2007). *The minimum description length principle*. Cambridge, MA: MIT Press.

Hoff, P. D., Raftery, A. E., & Handcock, M. S. (2002). Latent space approaches to social network analysis. *Journal of the american Statistical association*, *97*(460), 1090-1098.

Howell, R. D., Breivik, E., & Wilcox, J. B. (2007). Reconsidering formative measurement. *Psychological Methods*, *12*(2), 205–218. doi:10.1037/1082-989X.12.2.205

Ichikawa, M., & Konishi, S. (1999). Model evaluation and information criteria in covariance structure analysis. *British Journal of Mathematical and Statistical Psychology*, *52*(2), 285–302. doi:10.1348/000711099159116

Kendler KS, Prescott CA, Myers J, & Neale MC. (2003). The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Archives of General Psychiatry*, *60*(9), 929–937. doi:10.1001/archpsyc.60.9.929

Kroonenberg, P. m., & van Dam, M. (1997). Dynamics of behaviour in the strange situation: A structural... *British Journal of Psychology*, *88*(2), 311.

Lee, N., Cadogan, J., & Chamberlain, L. (2013). The MIMIC model and formative variables: Problems and solutions. *AMS Review, 1*(3), 3–17.

Myung, J. I., Navarro, D. J., & Pitt, M. A. (2006). Model selection by normalized maximum likelihood. *Journal of Mathematical Psychology, 50*, 167-179.

Rissanen, J. (1996). Fisher information and stochastic complexity. *IEEE Transactions on Information Theory*, *42*(1), 40–47.

Rissanen, J. (2001). Strong optimality of the normalized ML models as universal codes and information in data. *IEEE Transactions on Information Theory, 47*, 1712-1717.

Ryff, C. D., & Almeida, D. M. (2010). *National Survey of Midlife in the United States (MIDUS II): Daily Stress Project*, *2004-2009. ICPSR26841-v1*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. http://doi.org/10.3886/ICPSR26841.v1

Sampson, S. F. (1968). A Novitiate in a Period of Change: An Experimental and Case Study of Relationships. Ph.D. Thesis; Cornell University, Department of Sociology.

Shapiro, A. (1982). Rank-reducibility of a symmetric matrix and sampling theory of minimum trace factor analysis. *Psychometrika*, *47*, 187-199.

Tipping, M. E., & Bishop, C. M. (1999). Probabilistic principal component analysis. *Journal of the Royal Statistical Society: Series B*, *61*(3), 611–622. doi:10.1111/1467-9868.00196

Tumminello, M., Lillo, F., & Mantegna, R. N. (2007). Kullback-Leibler distance as a measure of information filtered from multivariate data. *Physical Review E*, 76, 031123. doi:10.1103/PhysRevE.76.031123

Van Der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, *113*(4), 842–861. http://doi.org/10.1037/0033-295X.113.4.842

Wang, J., Engelhard, G., & Lu, Z. (2014). Clarifying the conceptualization of indicators within different models. *Measurement: Interdisciplinary Research and Perspectives*, *12*(4), 155–159. doi:10.1080/15366367.2014.981124

West, S. G., & Grimm, K. J. (2014). Causal indicator models: unresolved issues of construction and evaluation. *Measurement: Interdisciplinary Research and Perspectives*, *12*(4), 160–164. doi:10.1080/15366367.2014.980108

Wood, D., Gardner, M. H., & Harms, D, P. (2015). How functionalist and process approaches to behavior can explain trait covariation. *Psychological Review*, *122*(1), 84–111. doi:10.1037/a0038423

Figure 1

[See attached jpeg]

Figure 1. Diagrams of the (a) reflective and (b) formative latent variable models. The x_i variables are proposed as indicators in both frameworks, though their content could be drawn from either structuralist or functionalist theory. λ_i and γ_i are the item loadings of reflective and formative constructs, respectively, on the latent variable η. ε_is are item variances, y_is are endogenous variables predicted by the latent trait, and ζ is a disturbance term.

Appendix A

Consider the case where observations are distributed according to a multivariate normal distribution with covariance matrix Σ. Assume that the observed data, with sample covariance matrix S, is represented using a saturated factor analytic model (e.g., minimum trace or minimum rank factor analysis; Shapiro, 1982). That is:

$$S=\Lambda\Lambda'+\Psi=(S-\Psi)+\Psi$$

where Λ is a matrix of loadings and Ψ is a diagonal matrix of residual variances. Then, following Tumminello, Lillo, and Mantegna (2007), it can be shown that the expected information lost by using the latent variables alone to describe the data is approximated by:

$$\frac{1}{2}\left[ln\left(\frac{|\Sigma-\Psi|}{|\Sigma|}\right)+tr([\Sigma-\Psi]^{-1}\Sigma)\right]+C$$

where $|\Sigma|$ is the determinant of Σ, $tr(\Sigma)$ is its trace, and C is a quantity depending on the number of persons and variables constituting the observations (cf. the maximum likelihood structural equation model fitting function; e.g., Ichikawa & Konishi, 2010). As the residual variances in Ψ go to zero, the logarithm term decreases to zero and trace term reduces to the number of variables. The information cost of using the latent variables to approximate the observed variables then reaches a minimum reflecting sampling variation. As the residual variances in Ψ increase, the information cost increases.