Annu. Rev. Psychol. 2002. 53:605–34

LATENT VARIABLES IN PSYCHOLOGY AND THE SOCIAL SCIENCES

Kenneth A. Bollen
Odum Institute for Research in Social Science, CB 3210 Hamilton, Department of Sociology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3210; e-mail: bollen@unc.edu

Key Words unmeasured variables, unobserved variables, residuals, constructs, concepts, true scores

■ **Abstract** The paper discusses the use of latent variables in psychology and social science research. Local independence, expected value true scores, and nondeterministic functions of observed variables are three types of definitions for latent variables. These definitions are reviewed and an alternative "sample realizations" definition is presented. Another section briefly describes identification, latent variable indeterminancy, and other properties common to models with latent variables. The paper then reviews the role of latent variables in multiple regression, probit and logistic regression, factor analysis, latent curve models, item response theory, latent class analysis, and structural equation models. Though these application areas are diverse, the paper highlights the similarities as well as the differences in the manner in which the latent variables are defined and used. It concludes with an evaluation of the different definitions of latent variables and their properties.

CONTENTS

0084-6570/02/0201-0605$14.00

INTRODUCTION

It is impossible to date the first use of latent variables. The idea that observable phenomena are influenced by underlying and unobserved causes is at least as old as religion, where unseen forces affect real world events. In the more secular sphere of everyday living, latent variables find wide application. From the response to "how are you feeling today?" to the description of a worker as "efficient" or a student as "bright," such abstract concepts elude direct measurement. What these examples illustrate is the common practice among humans to explain, to understand, and to sometimes predict events based on the role of concepts that are not directly observable. However, these more metaphysical, and everyday uses of unobserved forces depart from the use of latent variables in psychology and the social sciences. The scientific use of latent variables places a premium on designing research to test hypotheses about latent variables and having the ability to falsify hypotheses about them. In addition, latent variables provide a degree of abstraction that permits us to describe relations among a class of events or variables that share something in common, rather than making highly concrete statements restricted to the relation between more specific, seemingly idiosyncratic variables. In other words, latent variables permit us to generalize relationships.

Psychology has had its critics of latent variables, with Skinner (1976) being a well-known modern one. However, latent variables have been so useful in science that they pervade virtually all fields (see Glymour et al. 1987, pp. 22–26). Psychology and the social sciences are no exceptions. Although latent variables are part of numerous statistical and data analyses models, we do not have a single general definition of a latent variable that would include these diverse applications. Rather we have definitions of latent variables that are closely tied to specific statistical models and few systematic comparisons of these different definitions and the implications of the differences. Furthermore, the common problems that accompany the use of latent variables are obscured by the diverse definitions, each of which is tied to a limited number of applications.

Given the frequent appearance of latent variables in psychology and social science, it is surprising that so little work in these areas has focused on their nature. Borsboom et al. (2001), Edwards & Bagozzi (2000), Hägglund (2001), and Sobel (1994) are exceptions, but overall, my literature search concurs with Borsboom et al.'s conclusion that ". . . the theoretical status of the latent variable as it appears in models for psychological measurement has not received a thorough and general analysis as yet."

This paper aims to contribute to the discussion of latent variables in psychology and the social sciences. More specifically, the goals of the paper are (*a*) to review the major definitions of latent variables in psychology and the social sciences, (*b*) to formalize an intuitive and general definition of latent variables, (*c*) to examine latent variables in common statistical models in light of these definitions, and (*d*) to discuss issues that emerge when using latent variables.

I begin by reviewing several common ways of defining latent variables and introduce a "sample realizations" definition that is based on an intuitive notion of latent variables. With these definitions in hand, the next section discusses common properties and issues that arise when employing latent variables. Following this is a discussion of latent variables in a variety of statistical models including multiple regression, limited dependent-variable models (e.g., logistic and probit regressions), factor analysis, latent curve models, item response theory, latent class analysis, and structural equation models. The concluding section reviews the major findings from this review and highlights issues common to the use of latent variables.

DEFINITIONS OF LATENT VARIABLES

Unmeasured variables, factors, unobserved variables, constructs, or true scores are just a few of the terms that researchers use to refer to variables in the model that are not present in the data set. Many definitions of latent variables appear implicitly or explicitly. We can distinguish between nonformal and formal definitions. The next subsection briefly reviews several nonformal definitions. Four subsections that follow will present more formal definitions of latent variables: (*a*) local independence, (*b*) expected value, (*c*) nondeterministic function of observed variables, and (*d*) sample realization.[1] Next, I apply them to statistical models in psychology and the other social sciences in which latent or unobserved variables appear. This enables us to assess the applicability of these definitions across a range of areas.

Nonformal Definitions

One common set of definitions of latent variables considers them as "hypothetical variables." For instance, Harman (1960, p. 12) refers to factors as "hypothetical constructs." Similarly, Nunnally (1978, p. 96) defines a construct as something that scientists put together out of their imaginations (see also Bartlett 1937, p. 97). From this perspective, a property such as self-esteem is not real, but a hypothetical variable that comes from the mind of the researcher. This perspective contrasts with the Platonic view of latent variables in which the latent variables are seen as real (Sutcliffe 1965). Loevinger (1957, p. 642) makes the distinction between traits and constructs: "Traits exist in people; constructs (here usually about traits) exist in the minds and magazines of psychologists." Similarly, Edwards & Bagozzi (2000,

[1]These do not exhaust the formal definitions of latent variables. For example, Shafer (1996, pp. 352–56) uses probability trees and graph theory to briefly discuss latent variables. Similarly Pearl (2000) defines latent variables using graph theory. However, the definitions included here are among the most common formal definitions. Some ideas in the graph theory definitions are closely related to the local independence definitions of latent variables. It is too early to determine the impact of these graph theory–based definitions.

p. 156–57) view constructs as not real but as attempts to measure real phenomena. In their view the construct of self-esteem is not real, but there are real phenomena (or traits) to which researchers apply this term and construct.[2]

Another common definition type treats latent variables as impossible to measure, as unobservable or unmeasurable. Jöreskog & Sörbom (1979, p. 105) state that "latent variables . . . cannot be directly measured." Similarly the *Penguin Dictionary of Economics* (Bannock et al. 1998) defines a latent variable as "a variable in regression analysis which is, in principle, unmeasureable." These definitions presume knowledge that it is impossible to measure a latent variable. In a sense this presupposes that the researcher is able to know the future and that in that future there will be no innovations that will permit direct measurement of the latent variable. Thus, using this definition we would view self-esteem as not now directly measurable or measurable in the future. One difficulty with this definition is the assumption that we know the future and the impossibility of measuring a variable. Unforeseen technological or conceptual developments can occur that might make possible the measurement of variables that previously were treated as unmeasurable. A latent variable as unmeasurable definition does not permit this possibility.

A third type of informal definition defines latent variables as a data reduction device. Harman (1960, p. 5) says that ". . . a principle objective of factor analysis is to attain a parsimonious description of observed data." Thus, the latent variable or factor is a convenient means of summarizing a number of variables in many fewer factors. This definition gives primacy to the descriptive function of latent variables. It does not give much attention to latent variables that researchers define prior to analyzing the data or to the use of statistical procedures that test the implications of latent variable models. With this data reduction definition self-esteem is a term we might assign to a factor to summarize a group of items that "load" on a factor. The term is a shorthand expression for an underlying variable that helps explain the association between two or more variables.

It is possible to combine definitions. For instance, MacCallum & Austin (2000) state that "latent variables are hypothetical constructs that cannot be directly measured." Their definition combines the hypothetical and unmeasurable definitions. Taken individually or combined, these informal definitions do not capture all the ways in which researchers view latent variables. They appear best suited to exploratory analyses in which the nature of the latent variables and their relationships to observed variables is not specified in advance. Furthermore, these definitions are not based on formal definitions about the properties of the latent variables and do not provide technical assumptions about them.

[2]Pursuing this distinction between real and hypothetical variables leads to a metaphysical dilemma of deciding when something is real. Defining latent variables as only hypothetical narrows the use of the concept of latent variables and raises metaphysical debates on the meaning of "real variables." It seems preferable to leave the real or hypothetical nature of latent variables as an open question that may well be unanswerable.

Local Independence Definition

The "local independence" definition of a latent variable is one of the most common and popular ways to define a latent variable (Lord 1953, Lazarsfeld 1959, McDonald 1981, Bartholomew 1987, Hambleton et al. 1991). The key idea is that there are one or more latent variables that create the association between observed variables, and when the latent variables are held constant, the observed variables are independent. More formally,

$$P[Y_1, Y_2, \ldots, Y_K] = P[Y_1|\eta]P[Y_2|\eta]\cdots P[Y_K|\eta] \qquad 1.$$

where Y_1, Y_2, ..., Y_K are random observed variables, η is a vector of latent variables, $P[Y_1, Y_2, \ldots, Y_K]$ is the joint probability of the observed variables, and $P[Y_1|\eta]P[Y_2|\eta] \cdots P[Y_K|\eta]$ are the conditional probabilities. The joint probability of the observed variables equals the product of the conditional probabilities when the latent variables are responsible for the dependencies among the observed variables. In this definition we permit either continuous or discrete observed or latent variables in recognition of the variety of situations in which this definition of local independence applies. In a factor analysis, for instance, the latent and observed variables would be continuous; in item response theory continuous latent variables would appear in conjunction with discrete observed variables; in latent class analysis both observed and latent variables would be discrete.

McDonald (1981, 1996a) distinguished the above "strong" definition of local independence from a "weaker" form in which the linear association between variables is zero once the latent variables are held constant. An example of this weak form of the definition is

$$\rho_{Y_i Y_j \cdot \eta} = 0 \qquad 2.$$

for all i, j where $i \neq j$. $\rho_{Y_i Y_j \cdot \eta}$ is the partial correlation between two observed variables controlling for the vector of latent variables. If η contains the vector of latent variables underlying these observed variables, then this partial correlation will be zero once they are controlled. If the association remains, then we do not have the complete set of latent variables that underlie the data and we need to add more of them (Bartholomew 1987, p. 5). This is a weaker form of the local independence definition in that it refers to only the linear association between variables, whereas the stronger form of the definition refers to any dependence between the observed variables. Both forms of the local independence definitions define latent variables by their ability to completely explain the association of observed variables. Using this definition we could treat self-esteem as a latent variable if once it is held constant, there is no remaining dependence (or association) among the indicators that measure it. If dependence (association) remains, then we need to introduce additional latent variables or dimensions of self-esteem to capture it.

Several key implications of this definition are that it assumes (*a*) errors of measurement are independent (or uncorrelated), (*b*) observed variables or indicators have no direct or indirect effects on each other, (*c*) we have at least two observed variables, (d) each latent variable must have direct effects on one or more observed variables, and (e) the observed variables (indicators) do not directly affect the latent variable. As I illustrate below, these properties lead to counterintuitive elimination of some variables as latent variables.

Expected Value Definition

The expected value definition of a latent variable is most commonly associated with classical test theory (e.g., Lord & Novick 1968, Lumsden 1976, Jöreskog 1971). Here the term for the underlying variable is the "true score." The true score is equal to the expected value of the observed variable for a particular individual:

$$T_i \equiv \mathrm{E}(Y_i), \qquad 3.$$

where T_i is the value of the true score for the ith individual, $E(.)$ is the expected value, and Y_i is the random observed variable Y for the ith individual. This approach to defining a true score treats it as a value that would be obtained if we could perform a hypothetical experiment in which we could repeatedly observe Y_i for the ith individual without the responses being influenced by previous responses (Lord & Novick 1968, pp. 29–30). The mean of these infinitely replicated experiments would give us the true score value for that individual. Thus, rather than being defined by conditional independence among two or more observed variables, as in the preceding subsection, the expected value definition looks to the mean of the observed variable values for an individual as the true score. If we had an indicator of self-esteem for an individual, the true score on self-esteem would be the expected value of this measure under the hypothetical situation of repeatedly observing the indicator for the same individual where each trial would be independent of the others.

The equation for the observed random variable is

$$Y_i = T_i + E_i, \qquad 4.$$

where E_i is the error of measurement.

Several properties of the true score latent variable model are (a) its scale is defined by $E(Y_i)$; (b) the error of measurement, E_i, has a mean of zero and is uncorrelated with T_i; (c) the errors of measurement are uncorrelated for two different observed variables; (d) the true scores have direct effects on their corresponding observed variable; (e) the observed variables (indicators) do not directly affect the latent variable; and (f) two different observed variables have no direct or indirect effect on each other. As with the conditional independence definition, I argue that the true score latent variable model can lead to counterintuitive classifications of variables as latent or not.

Nondeterministic Function of Observed Variables Definition

Bentler (1982, p. 106) defines a latent variable as follows: "A variable in a linear structural equation system is a latent variable if the equations cannot be manipulated so as to express the variable as a function of manifest variables only." An interesting aspect of Bentler's definition is that it makes clear that we cannot use observed or manifest variables to exactly determine the latent variable. Although we might be able to manipulate the equations in which a latent variable appears, we cannot manipulate it to the point at which the latent variable is completely determined by the observed variables, that is, the latent variable is a nondeterministic function of the observed variables. In our hypothetical self-esteem example, self-esteem is a latent variable if we cannot manipulate its indicators to exactly express the self-esteem variable. We might be able to estimate or predict a value on the latent variable, but we would not be able to make an exact prediction based on its observed indicators.

This definition does not have the same exclusions as the local independence and expected value true score definitions of latent variables. It permits models with correlated errors of measurement and observed variables that directly or indirectly affect each other. The main restriction for this definition is that it is devised for linear structural equation systems and some latent variable models include nonlinear relations such as models with categorical observed variables. I illustrate below how the definition leads to disturbances being classified as latent variables in one model but not in another, whereas intuitively we would expect a consistent classification.

Sample Realization Definition

The "sample realization" definition that I provide is inspired by the simplest, intuitive understanding of a latent variable. Before giving more details on this definition, I provide a brief orientation of how I view latent variables. I present this orientation here rather than above because I do not assume that this perspective is shared by others who use different definitions of latent variables.

The starting point is the objects of study. The most common objects of study in psychology and the social sciences are individuals or groups. These objects have properties. Properties are characteristics of individuals or groups such as self-esteem, intelligence, cohesion, anxiety, etc. Theories hypothesize relations between these properties. For instance, we might theorize that intelligence promotes self-esteem. To test these ideas we build models. Models formalize the key elements in a theory. The individuals or groups are the objects (cases) in models. The variables in models represent the properties of objects and the model represents the relationships between the variables that are hypothesized in the theory. A model, for instance, could have a variable for self-esteem and another for intelligence, and the model would represent the hypothesized relation between them. The variables in the model are either manifest (observed) or latent (unobserved). Self-esteem and intelligence are both best represented as latent variables. More generally, our interest lies in the latent variables that are in models. The latent variables represent properties in a formal model, but they are not identical to these properties.

The definition of latent variables that I propose is a simple and inclusive definition of latent variables: A latent random (or nonrandom) variable is a random (or nonrandom) variable for which there is no sample realization for at least some observations in a given sample. In some ways this is not a new definition but is a formalization of a common idea that a latent variable is one for which there are no values. The definition permits the situation in which the random variable is latent (or missing) for some cases but not for others. In many situations a variable that is latent for any cases will be latent for all cases in a sample. The term "variable" in the definition refers to something that takes more than one value so that values that are constant across all cases are not included as variables.[3] Note also that the definition for random latent variables relies on the standard definition of a random variable.[4] The latent random variable differs from observed random variables in that for the observed random variable our sample contains realizations. If a random variable has realizations for some cases and not for others, then we can refer to it as latent (or missing) for those missing cases and an observed random variable for the other cases. Similarly, for nonrandom latent variables the variable takes more than one value, but if all or a subset of cases do not have sample realizations, then the variable is latent for those cases.

This definition of latent variables is rather minimalist and as such is more inclusive as to the variables considered as latent compared with the other definitions. For example, latent variables as defined by local independence are a special case of the sample realization definition, as are latent variables that conform to the expected value definition.

From the perspective of the sample realization definition all variables are latent until sample values of them are available. Of course, for many of the variables in the psychological and social sciences we do not have the option of directly observing such variables, so it will be latent for all cases in all samples. Our only option is to indirectly observe it through the sample values of an observed variable.

Another aspect of the definition is that it defines a variable as latent or not with respect to a particular sample. This implies that a variable could be latent in all, none, or just some samples. This permits the possibility that a variable is omitted in one sample but might be observed in another or it allows for the possibility that changes in techniques or advances in knowledge might allow us to measure variables previously treated as latent. For instance, before the invention of accurate thermometers, we could consider temperature a latent variable. But once

---

[3]Regression coefficient parameters, for example, would not be variables if the same parameter holds for all cases in a sample. Alternatively, for random coefficient models in which the regression parameters for the same variable differ across cases, the regression parameters would be a variable by this definition. For an example of the latter case, see the section on latent curve models.

[4]For example, "If *S* is a sample space with a probability measure and **x** is a real-valued function defined over the elements of *S*, then **x** is called a random variable" (Freund & Walpole 1987, p. 75).

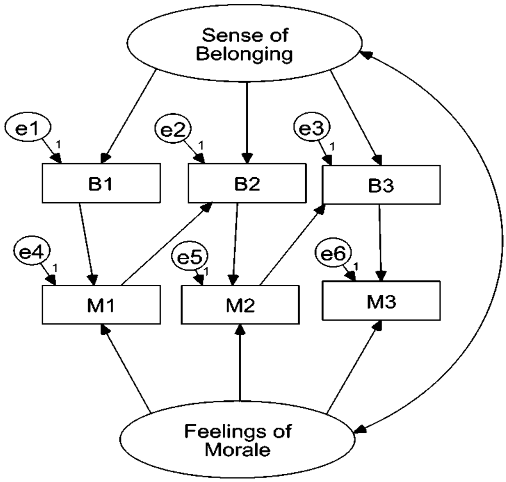

Figure 1 Model of “morale” and “sense of belonging” with response effects indicators (Bollen & Medrano 1998).

such thermometers are in use, their high accuracy permits us to treat their readings as a sample value of the previously latent variable.[5] Similarly to the degree that psychological and social measurement improve, we might reach the point where previously latent variables become observed variables.

The sample realizations definition permits models with correlated errors of measurement, observed variables that directly or indirectly influence each other, and many other nonstandard models. The key criterion is whether a variable has values for cases in a given sample.

Example

An example taken from Bollen & Medrano (1998) provides a means to further explore these definitions of latent variables. Figure 1 is a path diagram of a model with two unmeasured variables, “sense of belonging” and “feelings of morale,” enclosed in ovals. Observed random indicators of these variables are in boxes. The unique components of these indicators are also enclosed in ovals and point toward their respective measures. The straight single-headed arrows show the direct effect of the variable at the base of the arrow to the variable at the head of the arrow. The

[5]An argument could be made that contemporary thermometers are not perfect, so that the thermometer readings are not synonymous with temperature. However, the degree of measurement error in thermometer readings is miniscule compared with the error in social science measures. Considering thermometer readings as having negligible error is reasonable for the contrasts I wish to make.

curved two-headed arrow between belonging and morale signify the covariance between them. Three indicators of each unmeasured variable are part of the model. In addition, a response set effect is part of the model, where the response to one indicator partially drives the response to the next indicator on the questionnaire. For instance, the first indicator (M1) of feelings of morale follows the first indicator (B1) of sense of belonging. Therefore, the model shows a direct path from B1 to M1. Similarly, additional direct effects correspond to the order in which the indicators are given in the questionnaire.

Are morale and sense of belonging latent variables? The answer depends on the definition of latent variable that we use. In the nonformal definitions these variables would not be latent variables according to the data reduction definition because the primary purpose is not to come up with a descriptive reduction of the data. Rather, these indicators are created based on a theoretical definition of morale and belonging (see Bollen & Medrano 1998).

Whether these variables are inherently unmeasurable or impossible to directly measure is a problematic classification for the reasons described above: The definition presupposes knowledge that it will never be possible to directly measure these variables. Certainly we do not now have the technology or knowledge to do so, but we cannot say that it will never be possible. The last nonformal definition that describes latent variables as hypothetical variables requires a brief explanation before assessing whether morale and belonging are latent according to this definition. If we accept that latent variables are representations of properties of objects as I explained in the sample realization definition section, then the latent variables are not the same as the properties. In that sense they are hypothetical. However, if we push this hypothetical definition to apply to the property the latent variable represents, then the issue is less clear cut. Thus, it is possible that the properties of morale and belonging are real even though the latent variables that stand in for them in a model should not be reified.

Moving to the more formal definitions of latent variables, morale and belonging would not qualify as latent variables according to the local independence definition, because conditional on the values of morale and belonging, we cannot say that the six indicators are uncorrelated. Indeed, the model shows a direct relation between the indicators controlling for the unmeasured variables, thereby ruling out morale and belonging as latent variables. It is interesting to note that if there were no direct paths between the indicators shown in Figure 1, morale and belonging would conform to the local independence definition and thereby be latent variables. It seems counterintuitive to treat the variables as latent or not depending on whether the response set effects are in the model.

Morale and belonging are not latent variables when we apply the expected value definition. There are at least two problems that rule out these variables. One is that the expected value of all of the indicators except B1 would include a term that corresponds to the response set effect from the preceding variable. This would not be captured by morale and belonging. The second problem is that if each indicator has a unique component that is part of the error terms, then this unique component would contribute to the true score but would not be part of morale or belonging.

The nondeterministic function definition would classify morale and belonging as latent because we cannot write each latent variable as a deterministic function of the observed variable. Similarly, the sample realization definition would classify them as latent variables because we do not have sample realizations of these random variables. We can only indirectly observe them through their indicators.

As this example illustrates, the definition makes a difference in whether we would consider a variable as latent or not.

PROPERTIES OF LATENT VARIABLES

In addition to contrasting definitions of latent variables, it is useful to compare some of the issues that often accompany latent variables. In this section I highlight several contrasts that can be gleaned from the literature. The first important distinction is that between *a posteriori* and *a priori* latent variables. These terms are not used in the literature, but they do capture a distinction that is discussed. The *a posteriori* latent variables are latent variables that a researcher derives from the data analysis. In contrast, *a priori* latent variables are hypothesized prior to an examination of the data. The common distinction between exploratory and confirmatory factor analysis (Jöreskog 1969) helps capture this distinction. In exploratory factor analysis, the factors are extracted from the data without specifying the number and pattern of loadings between the observed variables and the latent factor variables. In contrast, confirmatory factor analysis specifies the number, meaning, associations, and pattern of free parameters in the factor loading matrix before a researcher analyzes the data (Bollen 1989, Ch. 7). Historically, the local independence definition of latent variables is closely tied to *a posteriori* latent variables in that latent variables (factors) are extracted from a set of variables until the partial associations between the observed variables goes to zero. The researcher defines the factors as part of a data reduction exercise.

Latent class analysis and other latent variable approaches also are distinguishable in whether they derive the latent variables from the data as part of the analysis or whether they use the data to test prespecified hypotheses about the latent variables. In practice it is probably best to regard the *a posteriori* and *a priori* as two points on a continuum in which most applications fall between these extremes.

A second issue is whether the latent variable is continuous, categorical, or a hybrid that falls between these ideal types. The question of whether the latent variable has gradations of values helps determine its nature. We cannot answer this question from the observed indicators of the latent variable, because it is possible to have a continuous, categorical, or hybrid observed variable with either a continuous, categorical, or hybrid latent variable. For instance, is depression a continuous latent variable with numerous gradations or are people either depressed or not, making it a categorical variable? Or should antisocial behavior be a dichotomy or a continuous variable with a floor of zero? Empirical means cannot always distinguish the nature of the latent variable from the empirical nature of the indicators (e.g., Bartholowmew 1987, Molenaar & von Eye 1994, Borsboom et al. 2001).

The third issue is the identification of the parameters associated with the latent variable in a model. Model identification asks whether it is possible to find unique values for the parameters that are in a model (Wiley 1973; Bollen 1989, Ch. 7; Davis 1993). Failure to achieve identification means that the factor loading or variance of a latent variable might not be unique and that we cannot tell the false from the true parameter values even if we have population data. Identification of latent variables are resolved differently, depending on the latent variable and the type of model, but usually it involves some minimal number of indicators or some constraints on the variance of the latent variables. A necessary condition for identification is that each latent variable must be assigned a scale. Though identification issues are present in simultaneous equations that ignore measurement error, the issues of identification in latent variable models raise additional complications.

Another issue is latent variable indeterminancy. This issue is well studied and debated in the factor analysis literature (e.g., Bartholowmew 1987, 1996; Green 1976; Guttman 1955; Maruan 1996a; McDonald 1996a; McDonald and Mulaik 1979; Mulaik 1996; Schönemann 1996; Steiger 1979, 1996a), but attempting to estimate latent variables from the observed variables is common across applications of latent variable models. Resolution of this indeterminancy is theoretically possible under certain conditions. Three conditions that can affect indeterminancy are (*a*) when the sample size (N) goes to infinity, (*b*) when the number of observed variables goes to infinity, and (*c*) when the squared multiple correlation for the latent variable goes to one and the predictors are observed variables. Of course, it would be highly unusual for one or more of these conditions to hold exactly, but it is possible for a condition to hold approximately, thereby approximately removing the indeterminancy. In the sections on statistical models we illustrate how these conditions can nearly remove the indeterminancy of the latent variable.

A final issue is that of whether the indicators of a latent variable are *causal indicators* or *effect indicators* (Blalock 1964, pp. 162–69; Bollen 1984; Bollen & Lennox 1991; Edwards & Bagozzi 2000). Causal (formative) indicators are observed variables that directly affect their latent variable. Examples include using time spent with friends, time spent with family, and time spent with coworkers as indicators of the latent variable of time spent in social interaction. Time spent watching violent television programs, time spent watching violent movies, and time spent playing violent video games would be causal indicators of exposure to media violence. Effect (reflective) indicators are observed variables that are effects of latent variables. Test scores on several tests of quantitative reasoning would be effect indicators of the latent variable of quantitative reasoning. Degree of agreement with questions about self-worth would be effect indicators of the latent variable of self-esteem. Nearly all measurement in psychology and the other social sciences assumes effect indicators. Factor analysis, reliability tests, and latent class analysis are examples of techniques that assume effect indicators. However, there are situations in which indicators are more realistically thought of as causes of the latent variable rather than the reverse. Tests for causal versus effect indicators have recently become available (Bollen & Ting 2000), but most empirical research implicitly assumes effect indicators. Incorrectly specifying indicators as causal or effect indicators leads to a misspecified model and holds the potential for inconsistent parameter estimates and misleading conclusions (Bollen & Lennox 1991).

In the next sections I illustrate how these varying definitions and properties apply to common statistical models.

LATENT VARIABLES IN STATISTICAL MODELS

Regression Disturbances as Latent Variables

Anyone who teaches courses on factor analysis or structural equation models is likely to have encountered skepticism when the idea of latent or unobserved variables is first mentioned. The reaction might be that such variables are too mystical or are something we should refrain from using. What is not fully appreciated is that it is quite likely that they have already been using unobserved, latent, or underlying variables in the other statistical procedures they have learned. To illustrate this, a convenient starting point is a multiple regression equation:

$$Y_i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_K X_{iK} + \varepsilon_i, \qquad 5.$$

where i indexes cases and runs from $i = 1, 2, \ldots, N$; Y_i is the value of the dependent observed random variable for the ith case; α is the equation intercept; β_k is the regression coefficient that corresponds to the X_{ik} variable where $k = 1, 2, \ldots, K$; and ε_i is the disturbance for the ith case.

It is of interest to examine the definitions of the disturbance term, ε_i. Some authors describe ε_i as a random variable that has three components: (*a*) an inherent, unpredictable random component present in virtually all outcomes, (*b*) a component that consists of a large number of omitted variables that influence Y_i, and (*c*) random measurement error in Y_i (e.g., Johnston 1984, pp. 14–15; Maddala 1988, p. 32). Other authors would add a fourth nonrandom component such as would occur if a researcher assumes a linear relation when a curvilinear one is more appropriate (e.g., Hanushek & Jackson 1977, pp. 12–13; Weisberg 1980, p. 6). Assuming that the nonrandom error is negligible, we can write the regression disturbance as

$$\varepsilon_i = \varepsilon_{ri} + \varepsilon_{oi} + \varepsilon_{mi}, \qquad 6.$$

where ε_{ri} is the inherently random component of the disturbance, ε_{oi} is a collection of the random omitted variables that influence Y_i, and ε_{mi} consists of random measurement error in measuring Y_i. ε_i and each of its components are unobserved variables that explain the discrepancy between Y_i and its predicted values based on the explanatory variables. The regression disturbance indicates a phenomenon in which the unobserved variable is a composite function of two or more latent variables rather than being a single component. In practice, researchers ignore the components of the regression disturbance and treat it as a unitary term, but this is not always the case.

If we consider that the analysis of variance and the analysis of covariance are special cases of multiple regression that also have disturbances, we readily see that much of psychology and the social sciences routinely use such unobserved or latent variables in their statistical modeling. Hence, to purge our models of unobservable or latent variables would require that we eliminate virtually all of the statistical techniques common in the social sciences.

Though the previous paragraphs use the term latent variable to describe the disturbance, not all of the definitions would include ε_i as a latent variable. The local independence definition presupposes at least two observed variables that depend on the latent variable. In multiple regression ε_i influences only Y_i. As such, ε_i would not qualify as a latent variable. The disturbance, ε_i, would also fail to satisfy the expected value definition of a true score (latent variable). By assumption, $E(\varepsilon_i)$ is zero for all cases, unlike the situation in which the expected value of an observed variable would take different values for different cases in the sample. More importantly, the expected value true score definition requires the expected value of an observed variable, whereas ε_i is unobserved.

According to Bentler's (1982, p. 107) definition, the disturbance would not be a latent variable because at the population level we can write $\varepsilon_i = Y_i - (\alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_{\mathrm{K}} X_{i\mathrm{K}})$. Thus, the disturbance is a function of observed variables only and hence does not satisfy Bentler's definition of a latent variable.[6] Alternatively, if we consider the components in Equation 6, each component of the disturbance would be a latent variable according to Bentler's definition, even though their sum, ε_i, would not be.

The sample realizations definition would qualify ε_i as a latent variable in that we do not have sample realizations in our sample data. We can estimate it as discussed below, but the estimates are not direct realizations of the random disturbance.

Viewing the disturbance, ε_i, as a latent variable provides the opportunity to introduce two issues common to all latent variables: identification issues and estimating values of the latent variable. Consider identification first. Every latent variable must be assigned a scale and a mean. Neither of these are inherent to a variable but instead are a matter of consensus among those working in an area. In the case of the disturbance in multiple regression, the disturbance is implicitly scaled to have the same units as the dependent variable, Y_i. This follows because the implicit

[6]In a sample we have $\hat{\varepsilon}_i$, the sample residual, and do not have the population disturbance, ε_i. From the perspective of the sample, ε_i is a latent variable because we cannot determine its exact value without the population regression coefficients (β_ks). Bentler (1982) does not make this distinction and considers the disturbance not to be a latent variable.

coefficient for ε_i is 1. Thus, a one-unit shift in ε_i leads to a one-unit shift in Y_i, holding constant the Xs.[7] The disturbance metric matches that of Y_i. The mean of ε_i is set to zero ($E(\varepsilon_i) = 0$). If we failed to make these assumptions, the multiple regression model would be underidentified and we would not be able to find unique values for at least some of the regression parameters. Even with these scaling assumptions, a multiple regression model is not identified. To identify it, we make another assumption about the latent variable, ε_i. We assume that the disturbance is uncorrelated with the Xs. Thus, using the sample realizations definition of a latent variable, the most widely used statistical procedure in the social and behavioral sciences, makes use of a latent variable called a disturbance and makes a number of assumptions about its behavior (coefficient of 1, $E(\varepsilon_i) = 0$, $COV(X_{ik}, \varepsilon_i) = 0$ where $i = 1, 2, 3, \ldots N$, $k = 1, 2, \ldots, K$).

A second issue that commonly accompanies the use of latent variables is attempts to estimate the values of the latent variable by using weighted combinations of the observed variables. This is linked to the issue of latent variable indeterminancy. In the case of multiple regression, "residuals" is a common name for the estimate of the disturbance, ε_i. The most widely used estimate of the disturbance latent variable is

$$\hat{\varepsilon}_i = Y_i - (\hat{\alpha} + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \cdots + \hat{\beta}_K X_{iK}), \qquad 7.$$

where $\hat{\varepsilon}_i$ contains the estimates of ε_i, $\hat{\alpha}$ is the ordinary least squares intercept estimator, and $\hat{\beta}_k$ is the ordinary least squares estimator of the regression coefficients. It is important to remember that the sample residuals, $\hat{\varepsilon}_i$ are not the same as the latent disturbances, ε_i. Unless $\hat{\alpha}$ and $\hat{\beta}_k$ match their corresponding population parameters, the sample residuals will not equal the population disturbances. As I noted above, one condition that sometimes removes latent variable indeterminancy is when the sample size, N, goes to infinity. In the above regression model, as $N \to \infty$, $\hat{\alpha} \to \alpha$ & $\hat{\beta}_k \to \beta_k$ and $\hat{\varepsilon}_i \to \varepsilon_i$, and the indeterminancy is removed. In practice, we have finite sample sizes so that at least some indeterminancy in the values of the disturbances are present. Furthermore, we generally do not have the information that would permit separate estimation of the three components of ε_i described above (ε_{ri}, ε_{oi}, ε_{mi}). Their sum is estimated in $\hat{\varepsilon}_i$, but their components remain indeterminant in a regression model even if $N \to \infty$.

Latent Variables in Limited Dependent-Variable Models

Multiple regression assumes that the dependent variable is continuous or nearly so. Categorical dependent variables are common in the social and psychological sciences and thus fall short of this assumption. Logistic and probit regression procedures permit noncontinuous dependent variables. *Limited dependent-variable models* is another term that refers to such models with categorical or censored

[7]Below, we introduce the assumption that ε_i is uncorrelated with the Xs. Under this assumption, we need not also assume that the Xs are held constant.

dependent variables (Maddala 1983, Long 1997). These models do not eliminate the need for latent variables, and from one perspective they make further use of latent variables than does the usual multiple regression. A convenient representation for limited dependent-variable models makes use of an equation that appears quite similar to a multiple regression,

$$Y_i^* = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_K X_{iK} + \varepsilon_i, \qquad 8.$$

where we can define all variables the same as the preceding multiple regression model except for the new symbol, Y_i^*. Y_i^* is a continuous unobserved variable that has a linear relation to the explanatory variables X_{ik}. The continuous Y_i^* is related to the categorical observed dependent variable. The nature of the relation depends on the nature of the observed categorical variable. For a dichotomous variable Y_i (= 1 or 0), the relation is

$$Y_i = \left\{ \begin{array}{ll} 1 & \text{if } Y_i^* > 0 \\ 0 & \text{if } Y_i^* \leq 0 \end{array} \right\}. \qquad 9.$$

This equation presents a threshold model where when Y_i^* exceeds 0, the dichotomous variable is one and when Y_i^* is at or below zero, the dichotomous variable is zero. The model assumes that underlying the dichotomous variable is a continuous variable that determines the category of the observed dichotomous variable. To illustrate, suppose the dichotomous variable asks whether a respondent agrees or disagrees with the statement, "I feel that I am as good as others." Though respondents differ widely in their degree of agreement or disagreement, they are left with only two options, agree or disagree. Equation 9 represents this as a threshold model in which once the unobserved degree of agreement passes a threshold of zero, the respondent will give an "agree" response. If they fall short of this threshold, the response will be "disagree." From one perspective, these threshold models are a correction of the crude way in which the original data were collected. From this viewpoint, the variables of interest are conceptualized as continuous, but the response format administered allows respondents to answer only in a restrictive, dichotomous scale.[8] In this formulation we have the disturbance, ε_i, as an unobserved variable as in multiple regression, but we also have Y_i^* as an underlying variable. The ε_i disturbance consists of the same components as in multiple regression ($\varepsilon_i = \varepsilon_{ri} + \varepsilon_{oi} + \varepsilon_{mi}$). Recall that in our discussion of multiple regression the local independence, expected value, and nondeterminant function of observed variables definitions would not classify ε_i as a latent variable, and the sample realization definition would treat it as latent. These same classifications hold for this limited dependent variable model with one exception. The nondeterminant function of observed variables definition would now classify ε_i as a latent variable. The reason is that even knowing the population parameters for

[8]It is possible to represent the dichotomous model without using a latent Y_i^* variable (see Long 1997 pp. 50–52).

all coefficients in Equation 8, we still cannot write ε_i as an exact function of observed variables because Y_i^* is unobserved.[9] The different definitions of latent variables would classify Y_i^* the same as ε_i: It is a latent variable according to the sample realization and nondeterminant function of observed variable definitions, but it is not according to the local independence and the expected value definitions.

To take things further we need to make further assumptions. We make the same assumptions about ε_i as we did in multiple regression [coefficient of 1, $E(\varepsilon_i) = 0$, $COV(X_{ik}, \varepsilon_i) = 0$]. These assumptions are sufficient to provide us with the mean of $Y_i^*[E(Y_i^*) = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_K X_{iK}]$, but the variance of $Y_i^*[VAR(Y_i^*) = VAR(\alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_K X_{iK}) + VAR(\varepsilon_i)]$ remains undetermined because we only can estimate the variance of $\alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_K X_{iK}$ and cannot estimate the variance of ε_i [$VAR(\varepsilon_i)$] and we need both to get the $VAR(Y_i^*)$. Strictly speaking, an assumption about the variance of ε_i would be sufficient to identify the variance of Y_i^*. In practice, the most common approach to defining the scale and variance of Y_i^* is to assume that the disturbance variable comes from a specific distribution. If ε_i comes from a standardized normal distribution [$\varepsilon_i \sim N(0,1)$], we have a dichotomous probit regression model. If we assume that ε_i comes from a standardized logistic distribution with a mean of zero and a variance of $\pi^2/3$, we are led to the dichotomous logistic regression model. Either of these assumptions provide the information required to identify the mean and scale for the two latent variables in the model, ε_i and Y_i^*.

This model is readily extended to ordinal outcome variables with more than two categories. Here we would maintain Equation 8, but the equation linking Y_i to Y_i^* becomes

$$Y_i = \left\{ \begin{array}{ll} 1 & \text{if } -\infty < Y_i^* \leq \tau_1 \\ 2 & \text{if } \tau_1 < Y_i^* \leq \tau_2 \\ 3 & \text{if } \tau_2 < Y_i^* \leq \tau_3 \\ \vdots & \\ C-1 & \text{if } \tau_{C-2} < Y_i^* \leq \tau_{C-1} \\ C & \text{if } \tau_{C-1} < Y_i^* \leq \infty \end{array} \right\}, \qquad 10.$$

where C refers to the total number of categories for the ordinal variable. The model is similar to the binary outcome model in that we have a nonlinear relation connecting the ordinal variable, Y_i, to Y_i^*, but it differs in that we have introduced "thresholds," τ_j, that are the cutpoints to determine into which category of the ordinal variable a case will fall. This model assumes that the value for the first threshold (τ_1) does not differ across the individuals in a sample. Similarly, the second threshold (τ_2) is constant over individuals, as are all other thresholds. Thus,

[9]Bentler's (1982) nondeterministic function of observed variables was proposed for linear structural equation models. The limited dependent variable models are nonlinear in the relation between the categorical dependent variable and the continuous underlying one, so his definition was not devised for such models.

the thresholds are similar to the regression coefficients in that they are population parameters that are the same for all cases in the sample.

As in the dichotomous case, we must ensure that it is possible to identify the mean and variances of the latent variables of ε_i and Y_i^*, but here we have the additional complication of identifying the thresholds. Like the dichotomous outcome model, assuming that ε_i is distributed as a standardized normal or standardized logistic variable will help to identify the means and variances of the latent variables, but now we also need to make assumptions about the thresholds. The two most common are to assume that τ_1 is zero or to assume that the equation intercept α is zero (see Long 1997, pp. 122–23). Either assumption in conjunction with the other distributional assumptions will identify the mean and variances of ε_i and Y_i^* and the thresholds, τ_j, in the model.

Limited dependent variable models such as these have been extended in a number of directions including censored variable models in which the continuous variable is directly observed for only part of its range and remains latent at some minimum or maximum value. Furthermore, multiple regression type models are sometimes combined with limited dependent-variable models such as in sample selectivity correction models (Heckman 1979, 1990; Greene 1997). From the perspective of this article the key aspect of these limited dependent-variable regression models is that they share with the usual multiple regression model the inclusion of unobserved variables. In addition to the unobserved disturbance random variable ε_i, there is a random underlying substantive variable (Y_i^*) that underlies the dichotomous, ordinal, or censored observed variable. The definitions of latent variables do not agree in classifying these variables as latent, and like multiple regression, each unobserved variable must be scaled and given a mean to permit identification of the model. Also similar to multiple regression, the explanatory variables in the model are observed variables.

A difference from multiple regression emerges when one tries to estimate the values of the unobserved variables. In multiple regression we could estimate the latent disturbance, ε_i, as $\hat{\varepsilon}_i = Y_i - (\hat{\alpha} + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \cdots + \hat{\beta}_K X_{iK})$. Attempting an analogous procedure for the limited dependent-variable model would lead to $\hat{\varepsilon}_i = Y_i^* - (\hat{\alpha} + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \cdots + \hat{\beta}_K X_{iK})$. Unfortunately we cannot calculate this quantity because Y_i^* is latent and we have no value to substitute for it. We can estimate the latent outcome variable, Y_i^*, as $\hat{Y}_i^* = (\hat{\alpha} + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \cdots + \hat{\beta}_K X_{iK})$ because we have X_{ik}s and estimates of β_k. The squared multiple correlation (R^2) calculated as described in McKelvey & Zavoina (1975, pp. 111–12) provides a measure of the "closeness" of this predicted latent variable to Y_i^*. The predicted version of the latent variable, $\hat{Y}_i^*$, should not be confused with the actual value of the latent variable. Just as $Y_i \neq \hat{Y}_i$ in a multiple regression with a continuous outcome, we have $Y_i^* \neq \hat{Y}_i^*$ in limited dependent-variable models because we cannot perfectly predict Y_i^*. As the R^2 goes to 1, $\hat{Y}_i^*$ goes to Y_i^* and this is an ideal condition under which the indeterminancy would disappear. Practice falls short of this ideal, so the indeterminancy of Y_i^* remains an issue.

Latent Variables in Factor Analysis

The factor analysis model is one of the first procedures psychologists would think of as a latent variable technique. Because factor analyses treat multiple indicators or observed variables at the same time, the factor analysis model is usually presented in a matrix form. However, I use a scalar form here because it helps point out the similarities and differences of the factor analysis model to the multiple regression and limited dependent-variable models of the prior sections (see section on general structural equation models for matrix expressions for factor analysis as part of the measurement model). Consider the equation for a single indicator from a factor analysis model,

$$Y_i = \lambda_0 + \lambda_1\xi_{i1} + \lambda_2\xi_{i2} + \cdots + \lambda_K\xi_{iK} + u_i, \qquad 11.$$

where Y_i is an observed variable or indicator for the ith case, λ_0 is an intercept term, λ_k is the "factor loading" that gives the impact of the kth factor on Y_i, ξ_{ik} is the kth factor for the ith case, and u_i is the "unique" variable or disturbance for the ith case. Factor analysis breaks the unique variable into two components,

$$u_i = s_i + e_i, \qquad 12.$$

where s_i is the specific component and e_i is the random measurement error, each of which is assumed to have a mean of zero and to be uncorrelated with each other and with the underlying factors (Harman 1960, Lawley & Maxwell 1971, Mulaik 1972). The specific component captures the systematic unique aspect of a variable that is uncorrelated with both the factors and with the random measurement error.

The factor analysis model (Equation 11) shares with multiple regression and the limited dependent-variable models the use of unobserved disturbances. The factor model, like multiple regression, has an observed dependent variable, whereas the limited dependent-variable model differs from both of these in its use of an underlying continuous dependent variable. Factor analysis departs from all prior models in its use of unobserved explanatory variables or "factors" as predictors of observed variables. The variables u_i, s_i, e_i, and ξ_{ik} are continuous latent random variables according to the sample realization and nondeterminant function definitions. Only ξ_{ik} are latent variables according to the local independence definition, provided that we have more than one indicator and that the correlation between these indicators goes to zero once the factors are controlled.

The expected value definition is more complicated when applied to this model. According to this definition, the expected value of Y_i would define the latent variable as

$$E(Y_i) = \lambda_0 + \lambda_1\mu_{\xi 1} + \lambda_2\mu_{\xi 2} + \cdots + \lambda_K\mu_{\xi_K}, \qquad 13.$$

where $\mu_{\xi k}$ is the mean of the ξ_{ik} factor. Thus, according to the expected value definition, the linear combination of the means of the factors would define a latent variable, but each separate ξ_k factor would not be a latent variable.

Each latent variable in the factor analysis model must be scaled. One way to scale each factor is to set one of the factor loadings from the factor to an observed variable to one. If for the same variable, we set the intercept to zero we also provide a mean for the factor. In the case in which this observed variable has only one factor influencing it, we get

$$Y_i = \xi_{i1} + u_i. \tag{14}$$

We can say that ξ_{i1} has the same scale and origin as Y_i in the sense that a one-unit change in ξ_{i1} leads to an expected change of one in Y_i and the latent and observed variable share the same mean (see Bollen 1989, pp. 307–11). An alternative scaling is to standardize each ξ_{ik} to a variance of one and a mean of zero. Other combinations are possible, but each factor must have a scale and an origin assigned. The unique component or disturbance u_i requires the same attention. Factor analysis models handle the scaling of u_i by giving it an implicit coefficient of one and setting its mean to zero.

Factor analysis provides a clear example of the distinction between the *a priori* and *a posteriori* latent variables raised above. In exploratory factor analysis the factors are *a posteriori* latent variables, that is, the factors are derived from the data rather than being defined before the analysis. Confirmatory factor analysis comes closer to the *a priori* latent variables because the factors and their pattern of loadings are determined prior to the data analysis. The *a posteriori* latent variables from exploratory factor analysis are closely associated with the tendency to see latent variables as hypothetical rather than real latent variables. This is easy to understand because the factors extracted in exploratory factor analysis are created by an algorithm and usually are only given "names" after extracted. This does not imply that the *a priori* latent variables in confirmatory factors are uniformly regarded as real.

As I discussed in the section on properties of latent variables, the indeterminancy of latent variables is well known in the factor analysis literature. However, under certain conditions the indeterminancy of a factor can in theory be removed. For instance, the squared correlation, ρ^2 (or "reliability coefficient") between the simple sum of indicators of a single factor and that factor is[10]

$$\rho^2 = \frac{\left(\sum_{j=1}^{J} \lambda_j\right)^2 VAR(\xi_1)}{\left(\sum_{j=1}^{J} \lambda_j\right)^2 VAR(\xi_1) + \sum_{j=1}^{J} VAR(u_j)}. \tag{15}$$

The j indexes the indicators of the latent ξ_1 factor, $j = 1, 2, \ldots, J$, and each indicator loads only on ξ_1 with a unique component that has a mean of zero and is uncorrelated with all other unique components for the other indicators. With some

[10]This formula is derivable from Bollen (1980, p. 378) when there are no correlated errors of measurement.

algebraic manipulations, I rewrite Equation 15 as

$$\rho^2 = \frac{1}{1 + \dfrac{\sum_{j=1}^{J} VAR(u_j)}{\left(\sum_{j=1}^{J} \lambda_j\right)^2 VAR(\xi_1)}}. \qquad 16.$$

This equation reveals that ρ^2 goes to one and the factor indeterminancy is removed when the second term in the denominator goes to zero. For instance, suppose that the latent factor and all indicators are standardized to a variance of one and that every indicator has a standardized factor loading of 0.7. Ten such indicators would result in a ρ^2 of 0.94 for their simple sum; 50 indicators would have a squared correlation of 0.99. Of course, the rate of growth in the squared correlation and hence in lessening indeterminancy depends on the magnitude of the factor loadings and errors in addition to the number of indicators, but this example illustrates how increasing the number of indicators ($K \to \infty$) of a single factor can reduce indeterminancy. See Piaggio (1931, 1933), Mulaik & McDonald (1978), and McDonald & Mulaik (1979) for further discussion of the relation between the number of indicators and indeterminancy of factor scores in exploratory factor analysis.

Latent Curve Models

Latent curve models apply to longitudinal data in which repeated measures are available for the same cases (e.g., Tucker 1958, Meredith & Tisak 1990, McArdle & Hamagami 1991, Willett & Sayer 1994). Though it is possible to formulate these models for categorical outcomes, I limit the discussion to continuous repeated measures. The equations for an unconditional latent curve model are

$$Y_{it} = \alpha_i + \beta_i \lambda_t + \varepsilon_{it}, \qquad 17.$$

$$\alpha_i = \mu_\alpha + \zeta_{\alpha i}, \qquad 18.$$

$$\beta_i = \mu_\beta + \zeta_{\beta i}, \qquad 19.$$

where $i = 1, 2, \ldots, N$ indexes individuals, $t, = 0, 1, \ldots, T$ indexes time, α_i is the intercept for the ith case, β_i is the slope of the trajectory for the ith case, $\lambda_t = 0, 1, \ldots, T$ is a time trend variable, ε_{it} is a disturbance for the itth observation, μ_α and μ_β are the means of the intercepts and slopes, and $\zeta_{\alpha i}$ and $\zeta_{\beta i}$ are disturbances. All disturbances are scaled by setting their means to zero and their coefficients to one in the equation in which they appear. By assumption, ε_{it} is uncorrelated with α_i, β_i, $\zeta_{\alpha i}$, and $\zeta_{\beta i}$.

The latent curve model departs from the others we have considered in that the random coefficients, α_i and β_i are unobserved variables. From the perspective of the local independence definition, these are latent variables as long as we have at least two waves of data for the Ys, though we generally require at least three waves of data to identify the model. Similarly, they are latent variables applying the other definitions. This is straightforward for these definitions except for the expected value one. To discuss the expected value definition further, consider the equation for Y_{i1},

$$Y_{i1} = \alpha_i + \varepsilon_{i1}. \tag{20}$$

Note the similarity of this equation to the true score Equation 4. If we could repeatedly observe Y_{i1} in the sense that I discussed under the expected value definition, then the expected value of Y_{i1} would be α_i, where α_i is a constant intercept for the *i*th case.[11] An analogous argument holds for β_i if we use the trick of taking difference scores,

$$Y_{i2} - Y_{i1} = (\alpha_i + \beta_i + \varepsilon_{i2}) - (\alpha_i + \varepsilon_{i1}) \tag{21}$$

$$= \beta_i + (\varepsilon_{i2} - \varepsilon_{i1}). \tag{22}$$

Using difference scores (Equation 22) appears to be in a form that conforms to the expected value definition of a true score or latent variable. However, one complication is that the errors of measurement for Equations 20 and 21 are correlated, and this violates one of the assumptions for the expected value definition of latent variables. Thus, these only partially satisfy the expected value definition of latent variables, and under the strict definition these would not be latent variables.

The disturbance terms, ε_{it}, $\zeta_{\alpha i}$, and $\zeta_{\beta i}$, are not latent variables according to two definitions. More specifically, only the nondeterministic function and sample realizations definitions would classify these as latent, whereas the local independence and expected value definitions would not. These disturbances and the random coefficients would all be *a priori* in that they are hypothesized prior to the data analysis.

Item Response Theory

Item response theory (IRT) refers to a collection of related techniques that have wide application in psychological measurement (see, e.g., Lord 1980, Hambleton & Swaminathan 1985). They are well suited to handle dichotomous or ordinal observed variables. Though there are many different models for IRT, a simple one-parameter logistic model for dichotomous variables can illustrate the key points with respect to latent variables. We can write this model as

$$P_j(\xi) = \frac{e^{(\xi - \kappa_j)}}{1 + e^{(\xi - \kappa_j)}}, \tag{23}$$

where ξ is the underlying "ability" variable, *e* is the mathematical constant, κ_j the item difficulty parameter, *j* indexes the item (or observed dichotomous variable), and $P_j(\xi)$ is the probability that an item *j* is correct at ability ξ. κ_j is the item

[11]This is different than taking the expected value of the random intercepts over all individuals. In this case the expected value over individuals would be μ_α.

difficulty parameter because the higher its value, the more difficult (i.e., the lower the probability) it is to get a "correct" response to an item. The probability of a correct response for two different items will differ even if the latent variable is at the same value if the item difficulty parameters differ. These models are similar to the limited dependent-variable model in that the observed outcome variable is categorical, but it departs from these models in having an unobserved determinant of the outcome.

Given the nonlinear function connecting the dichotomous item and the underlying variable, ξ would not qualify as a latent true score according to the expected value definition. ξ is not a deterministic function of the observed variables, so this definition of a latent variable would be satisfied with the qualification that the nondeterministic function definition was proposed for linear models. However, the local independence definition of a latent variable is key to IRT, so according to this definition, ξ would be a latent variable provided we have at least two items for the same latent variable. Similarly, the sample realization definition would be satisfied and ξ would be latent by this definition as well. Identifying this model requires that we scale and provide a mean for the latent variable. A common approach is to set the ξ variable's mean to zero and its variance to one (Hambleton et al. 1991, p. 42). Summing items provides a method to estimate the latent variable, but the issue of latent variable indeterminancy remains. Increasing the number of items that tap the unidimensional latent variable can increase the correlation between the sum of the items and the latent variable, but in practice some indeterminancy will persist. More complicated IRT models are available, but the classification of variables as latent or not would follow a similar pattern as that described above.

Latent Class Analysis

In all of the models reviewed the unobserved explanatory variables have been continuous even though the observed variables could be categorical or continuous. In this section I briefly present a model in which both the underlying variable and the observed variables are categorical variables. Lazarsfeld's *latent class model* (Lazarsfeld 1950, 1959; Anderson 1954, 1959; Lazarsfeld & Henry 1968; Goodman 1978; Langeheine & Rost 1988; Heinen 1996) has considerable generality, but to simplify the presentation I only consider a situation in which there are three dichotomous observed variables, X_1, X_2, and X_3, and one dichotomous underlying variable, ξ. The observed and underlying variable each have only two possible values, 0 or 1. The fundamental equation of latent structure analysis is

$$P(X_1 = c_1, X_2 = c_2, X_3 = c_3) = \sum_{c=0}^{1} P(\xi = c)P(X_1 = c_1|\xi = c) \times P(X_2 = c_2|\xi = c)P(X_3 = c_3|\xi = c), \quad 24.$$

where $P(.)$ refers to an unconditional probability, $P(.|.)$ is a conditional probability, c_1, c_2, c_3, and c refer to the value of 0 or 1 for X_1, X_2, X_3, or ξ. This equation says that the unconditional probability of a triplet set of values (c_1, c_2, c_3) for the three observed dichotomous variables (X_1, X_2, X_3) is equal to the sum over c of the unconditional probabilities of the latent variable being in the cth category times the conditional probabilities for each of the observed variables given that the underlying variable is in the cth category. Underlying this probability is the assumption of local independence. That is, any association between X_1, X_2, and X_3 is due to their common dependence on ξ. Within categories of ξ the observed dichotomous variables are independent.

It is not surprising to find that the latent class model conforms to the local independence definition of a latent variable so that using it, we can refer to ξ as a latent variable. The sample realization definition of latent variables would also treat ξ as a latent variable because there are only indicators of it but no direct observations. The nondeterministic function definition was intended for linear structural equation models, but in general we cannot write ξ as an exact function of the dichotomous observed variables so it would be latent. In contrast, the expected value definition would not classify ξ as latent.

Analogous to factor analysis, both *a posteriori* and *a priori* latent variables might appear in latent class analysis. It depends on whether the latent variables are hypothesized before or after the data analysis. Furthermore, indeterminancy of the latent class is an issue, and again, like factor analysis, the number of indicators with properties that conform to the model can lessen the degree of indeterminancy.

Structural Equation Models with Latent Variables

Structural equation models are widely used in psychology and the social sciences. In their most general form they include most of the models from the previous sections (Goldberger & Duncan 1973; Jöreskog 1977; Bentler & Weeks 1980; Muthén 1984; Bollen 1989, 1998; Muthén & Muthén 2001). A slight modification of the LISREL notation presents the model as[12]

$$\eta = \alpha_\eta + \mathbf{B}\eta + \Gamma\xi + \zeta \qquad 25.$$

$$\mathbf{Y} = \alpha_{\mathrm{Y}} + \Lambda_Y\eta + \varepsilon \qquad 26.$$

$$\mathbf{X} = \alpha_X + \Lambda_X\xi + \delta. \qquad 27.$$

Equation 25 is the latent variable model where η is a vector of latent endogenous variables with $\mathbf{B}$ a matrix of regression coefficients for the impact of the latent endogenous variables on each other, $\boldsymbol{\xi}$ is the vector of latent exogenous variables with $\boldsymbol{\Gamma}$ a matrix of regression coefficients for the latent exogenous variable's impact on the latent endogenous variables, $\boldsymbol{\alpha}_\eta$ is a vector of equation intercepts, and $\boldsymbol{\zeta}$ is the vector of latent disturbances that have a mean of zero and are uncorrelated with $\boldsymbol{\xi}$. Equations 26 and 27 are the measurement model equations in which the

[12]To simplify the discussion, I only present this model for the continuous latent and observed variables. For a discussion of extending this model to categorical variables, see the more recent references cited in the previous sentence.

former relates **Y**, a vector of observed variables, to $\boldsymbol{\eta}$ via a coefficient matrix of factor loadings, $\boldsymbol{\Lambda}_Y$, $\boldsymbol{\alpha}_Y$ is a vector of equation intercepts, and $\boldsymbol{\varepsilon}$ is a vector of unique components that have a mean of zero and are uncorrelated with $\boldsymbol{\eta}$, $\boldsymbol{\xi}$, and $\boldsymbol{\zeta}$. Equation 27 is similarly defined as the indicators for the $\boldsymbol{\xi}$ latent variables. Each equation alone can represent a factor analysis model (Bollen 1989, Ch. 7). As such our previous discussion of the classification of variables as latent or not according to the different definitions of latent variables carries over. Because $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ are classified as latent variables via these measurement models (Equations 26 and 27), the only unconsidered unmeasured variable is $\boldsymbol{\zeta}$, but the classification of this variable closely follows the previous discussion of disturbances as latent variables.

DISCUSSION AND CONCLUSIONS

One clear conclusion from this review is that whether we consider a variable latent or not depends on the definition we use. Table 1 summarizes the classifications of the major unmeasured variables that appear in the different statistical models reviewed in this paper. It classifies them as latent according to four definitions of latent variables: local independence, expected value, nondeterminant function, and sample realization definitions. Table 1 reveals that the most inclusive definition is the sample realization definition, closely followed by the nondeterminant function definition. The local independence is perhaps the most common definition of latent variables, yet Table 1 shows it to be fairly restrictive, only exceeded by the restrictiveness of the expected value definition. The other three definitions of latent

TABLE 1 Summary classification of latent variables in statistical models

	Definition			
Model	**Local independence**	**Expected value**	**Nondeterministic function**	**Sample realization**
Multiple regression ε_i	No	No	No	Yes
Limited dep. var. Y_i^*, ε_i	No	No	Yes	Yes
Factor analysis u_i, ξ_{ik}	u_i No, ξ_{ik} Yes[a]	No	Yes	Yes
Latent curve α_i, β_i	Yes[a]	Partially[b]	Yes	Yes
ε_{it}, $\zeta_{\alpha i}$, $\zeta_{\beta i}$	No	No	Yes	Yes
Item response theory ξ	Yes[a]	No	Yes	Yes
Latent class ξ	Yes[a]	No	Yes	Yes
Structural equations $\eta, \xi, \zeta, \varepsilon, \delta$	η, ξ sometimes $\zeta, \varepsilon, \delta$ no	No	Yes	Yes

[a]Assumes that error terms are uncorrelated.

[b]Errors for observed variables are correlated contrary to expected value definition, though they otherwise conform. See text.

variables are restrictive forms of the sample realization definition. For instance, if we use the sample realization definition and impose the restriction that the observed variables be independent once the latent variables are controlled, we are led to the local independence definition. The nondeterminant function definition is less restrictive than the local independence and expected value definitions. One of its limitations is that it was devised for linear structural equations, whereas there are other models in which latent variables appear that would not be covered by this definition.

Though the local independence and the expected value definitions are useful in some contexts, they lead to counterintuitive classifications of variables as latent or not. For instance, the local independence and expected value definitions do not classify disturbances as latent variables. In contrast, Arbuckle & Wothke (1999) refer to disturbances as latent variables. Griliches (1974, pp. 976–77) classifies disturbances as one of the three types of latent ("unobservables") variables.

A similar problem occurs in any factor analyses or structural equation models with correlated errors of measurement. The local independence definition would not be satisfied. An example is where we have a single factor with four indicators and correlated errors between the second and third indicators. If the correlated errors were absent, the factor would satisfy the local independence definition for a latent variable. Or I could replace the correlated errors with a single unmeasured variable, uncorrelated with the other factor, with factor loadings fixed to one for the second and third measures. Now both factors would qualify as latent. However, it seems counterintuitive to consider the underlying variable in the model not latent when there are correlated errors, but latent when the correlated errors are replaced with an additional factor.

Higher order factors would be excluded as latent variables with the expected value definition, but these hold an ambiguous status with the local independence definition. If we consider that the indicator variables are uncorrelated once we control for all first order factors, then the higher order factor seems not needed and its status as a latent variable is ambiguous with the local independence definition. This is true even though the first order factors from the same model satisfy the local independence definition when the errors are uncorrelated. Similarly, the unmeasured variable influenced by causal indicators would not be latent according to the local independence definition unless it had at least two effect indicators with uncorrelated errors of measurement. As in the case of disturbances, the sample realization definition would treat the underlying variables as latent in all of these examples.

What explains these different definitions of latent variables? Part of the explanation is that the definitions emerged from different statistical models. For instance, the expected value definition came out of the classical test theory, whereas the local independence definition has roots in latent class and factor analysis. The nondeterministic function of observed variables originates with linear factor analysis and structural equation models. The sample realization definition proposed here was the most inclusive definition because the only requirement is that there is not a sample realization of a variable for a case in a given sample. It is not based on any one of these statistical models but attempts to apply to all of them.

An advantage of the sample realization definition is that it helps make connections between underlying variables in a variety of models and applications. For instance, the distinction between *a posteriori* and *a priori* latent variables holds for all such variables. Issues of identification and scaling are common across these unmeasured variables, as is the problem of latent variable indeterminancy. We also need to decide on the direction of influence between the manifest and latent variables, that is, are there causal indicators or effect indicators? As this review reveals, most attention is directed toward effect indicators, but the sample realization definition holds for unmeasured variables whether there are causal or effect indicators.

In conclusion, there is no right or wrong definition of latent variables. It is more a question of finding the definition that is most useful and that corresponds to a common understanding of what should be considered latent variables. If we stick with the conventional dichotomy of variables being either latent or observed, several interesting questions are posed for the most restrictive definitions. For instance, the factors in a factor analysis model with correlated errors are not latent variables according to the local independence and expected value definitions. They certainly are not observed variables, so what types of variables are they if not latent or observed? Similarly how do we classify disturbances or errors? If we wish to add additional categories of variables beyond latent and observed, what do we gain by creating these new categories? Also, do we miss common properties of variables across these categories by giving them distinct names? These are questions that must be answered if we use the more restrictive definitions of latent variables.

ACKNOWLEDGMENTS

Partial support for this research comes from NIDA grant DA13148. I am grateful for the comments and help from the Carolina Structural Equation Modeling (CSEM) group, the SEMNET Listserv, Denny Borsboom, Susan Fiske, Jennifer Glanville, and Keith Markus. An earlier version of the paper was presented at the International Meeting of the Psychometric Society, July, 2001, Osaka, Japan.

LITERATURE CITED

Anderson TW. 1954. On the estimation of parameters in latent structure analysis. *Psychometrika* 19:1–10

Anderson TW. 1959. Some scaling methods and estimation procedures in the latent class model. In *Probability and Statistics*, ed. U Grenander, pp. 9–38. New York: Wiley

Arbuckle J, Wothke W. 1999. *Amos 4.0*. Chicago: Small Waters

Bannock G, Baxter RE, Davis E. 1998. *Penguin Dictionary of Economics*. New York: Penguin

Bartholowmew DJ. 1987. *Latent Variable Models and Factor Analysis*. London: Griffin

Bartholowmew DJ. 1996. Comment on: metaphor taken as math: indetermimacy in the factor model. *Multivariate Behav. Res.* 31:551–54

Bartlett MS. 1937. The statistical conception of mental factors. *J. Psychol.* 28:97–104

Bentler PM. 1982. Linear systems with multiple levels and types of latent variables. In *Systems Under Indirect Observation*, ed. KG Jöreskog, H Wold, pp. 101–30. Amsterdam: North-Holland

Bentler PM, Weeks DG. 1980. Linear structural equations with latent variables. *Psychometrika* 45:289–308

Blalock HM. 1964. *Causal Inferences in Nonexperimental Research*. Chapel Hill: Univ. NC Press

Bollen KA. 1980. Issues in the comparative measurement of political democracy. *Am. Sociol. Rev.* 45:370–90

Bollen KA. 1984. Internal consistency or no necessary relationship? *Qual. Quant.* 18:377–85

Bollen KA. 1989. *Structural Equations with Latent Variables*. New York: Wiley

Bollen KA. 1998. Structural equation models. In *Encyclopedia of Biostatistics*, ed. P Armitage, T Colton, pp. 4363–72. Sussex, UK: Wiley

Bollen KA, Lennox R. 1991. Conventional wisdom on measurement: a structural equation perspective. *Psychol. Bull.* 110:305–14

Bollen KA, Medrano JD. 1998. Who are the Spaniards?: nationalism and identification in Spain. *Soc. Forces* 77:587–621

Bollen KA, Ting K. 2000. A tetrad test for causal indicators. *Psychol. Methods* 5(1):3–22

Boorsboom D, Mellenbergh GJ, Heerden JV. 2001. Philosophy of science and psychometrics: Reflections on the theoretical status of the latent variable. Methodological Rep. 20011. Amsterdam: Univ. Amsterdam Dept. Psychol.

Davis W. 1993. The FC1 rule of identification for confirmatory factor analysis. *Sociol. Methods Res.* 21:403–37

Edwards JR, Bagozzi RP. 2000. On the nature and direction of relationships between constructs and measures. *Psychol. Methods* 5:155–74

Freund JE, Walpole RE. 1987. *Mathematical Statistics*. Englewood Cliffs, NJ: Prentice Hall

Glymour C, Scheines R, Spirtes P, Kelly K. 1987. *Discovering Causal Structure: Artificial Intelligence, Philosophy of Science, and Statistical Modeling*. New York: Academic

Goldberger A, Duncan OD. 1973. *Structural Equation Models in the Social Sciences*. New York: Academic

Goodman LA. 1978. *Analyzing Qualitative/Categorical Data: Log-Linear Models and Latent Structure Analysis*. Cambridge, MA: Abt Books

Green BF. 1976. On the factor scores controversy. *Psychometrika* 41:263–66

Greene WH. 1997. *Econometric Analysis*. Upper Saddle River, NJ: Prentice Hall

Griliches Z. 1974. Errors in variables and other unobservables. *Econometrica* 42:971–98

Guttman L. 1955. The determinacy of factor score matrices with implications for five other basic problems of factor theory. *Br. J. Stat. Psychol.* 8:65–82

Hägglund G. 2001. Milestones in the history of factor analysis. In *Structural Equation Modeling: Present and Future*, ed. R Cudeck, S du Toit, D Sörbom, pp. 11–38. Lincolnwood, IL: Scientific Software Int.

Hambleton RK, Swaminathan H. 1985. *Item Response Theory: Principles and Applications*. Boston: Kluwer

Hambleton RK, Swaminathan H, Rogers HJ. 1991. *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage

Hanushek EA, Jackson JE. 1977. *Statistical Methods for Social Scientists*. New York: Academic

Harman HH. 1960. *Modern Factor Analysis*. Chicago: Univ. Chicago Press

Heckman J. 1979. Sample selection bias as a specification bias. *Econometrica* 47:153–61

Heckman J. 1990. Varieties of selection bias. *Am. Econ. Rev.* 80:313–18

Heinen T. 1996. *Latent Class and Discrete Latent Trait Models*. Thousand Oaks, CA: Sage

Johnston J. 1984. *Econometric Methods*. New York: McGraw-Hill

Jöreskog KG. 1969. A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika* 34:183–202

Jöreskog KG. 1971. Statistical analysis of sets of congeneric tests. *Psychometrics* 36:109–33

Jöreskog KG. 1977. Structural equation models in the social sciences: specification, estimation, and testing. In *Applications of Statistics*, ed. RC Atkinson, DH Krantz, RD Luce, P Suppes, pp. 265–87. Amsterdam: North-Holland

Jöreskog KG, Sörbom D. 1979. *Advances in Factor Analysis and Structural Equation Models*. Cambridge, MA: Abt Books

Langeheine R, Rost J. 1988. *Latent Trait and Latent Class Models*. New York: Plenum

Lawley DN, Maxwell AE. 1971. *Factor analysis as a Statistical Method*. London: Butterworth

Lazarsfeld PF. 1950. The logical and mathematical foundations of latent structure analysis. In *The American Soldier: Studies in Social Psychology in World War II*, Vol. IV. *Measurement and Prediction*, ed. SA Stouffer, pp. 362–412. Princeton, NJ: Princeton Univ. Press

Lazarsfeld PF. 1959. Latent structure analysis. In *Psychology: A Study of Science*, ed. S Koch, pp. 476–543. New York: McGraw-Hill

Lazarsfeld PF, Henry NW. 1968. *Latent Structure Analysis*. Boston: Houghton Mifflin

Loevinger J. 1957. Objective tests as instruments of psychological theory. *Psychol. Rep.* 3:1–18

Long JS. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage

Lord FM. 1953. The relation of test score to the trait underlying the test. *Educ. Psychol. Meas.* 13:517–49

Lord FM. 1980. *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale, NJ: Erlbaum

Lord FM, Novick MR. 1968. *Statistical Theories of Mental Test Scores*. Reading, MA: Addison-Wesley

Lumsden J. 1976. Test theory. *Annu. Rev. Psychol.* 27:251–80

MacCallum RC, Austin JT. 2000. Applications of structural equation modeling in psychological research. *Annu. Rev. Psychol.* 51: 201–26

Maddala GS. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. New York: Macmillan

Maddala GS. 1988. *Introduction to Econometrics*. New York: Macmillan

Maraun MD. 1996a. Metaphor taken as math: indeterminancy in the factor analysis model. *Multivariate Behav. Res.* 31:517–38

Maraun MD. 1996b. The claims of factor analysis. *Multivariate Behav. Res.* 31:673–89

McArdle JJ, Hamagami F. 1991. Modeling incomplete longitudinal and cross-sectional data using latent growth structural models. In *Best Methods for the Analysis of Change*, ed. LM Collins, JC Horn, pp. 276–304. Washington, DC: Am. Psychol. Assoc.

McDonald RP. 1981. The dimensionality of tests and items. *Br. J. Math. Stat. Psychol.* 34:100–17

McDonald RP. 1996a. Latent traits and the possibility of motion. *Multivariate Behav. Res.* 31:593–601

McDonald RP. 1996b. Consensus emergens: a matter of interpretation. *Multivariate Behav. Res.* 31:663–72

McDonald RP, Mulaik SA. 1979. Determinacy of common factors: a non-technical review. *Psychol. Bull.* 86:297–306

McKelvey RD, Zavoina W. 1975. A statistical model for the analysis of ordinal level dependent variables. *J. Math. Sociol.* 4:103–20

Meredith W, Tisak J. 1990. Latent curve analysis. *Psychometrika* 55:107–22

Molenaar PCM, von Eye A. 1994. On the arbitrary nature of latent variables. See von Eye & Clogg 1994, pp. 226–42

Mulaik SA. 1972. *The Foundations of Factor Analysis*. New York: McGraw-Hill

Mulaik SA. 1996. On Maraun's deconstructing of factor indeterminancy with constructed factors. *Multivariate Behav. Res.* 31:579–92

Mulaik SA, McDonald RP. 1978. The effect of additional variables on factor indeterminancy in models with a single common factor. *Psychometrika* 43:177–92

Muthén B. 1984. A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika* 49:115–32

Muthén LK, Muthén B. 2001. *Mplus: Statistical Analysis with Latent Variables, User's Guide*. Los Angeles: Muthén & Muthén

Nunnally JC. 1978. *Psychometric Theory*. New York: McGraw-Hill

Pearl J. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge, UK: Cambridge Univ. Press

Piaggio HTH. 1931. The general factor in Spearman's theory of intelligence. *Nature* 127:56–57

Piaggio HTH. 1933. Three sets of conditions necessary for the existence of a *g* that is real and unique except in sign. *Br. J. Psychol.* 24: 88–105

Schönemann PH. 1996. The psychopathology of factor indeterminancy. *Multivariate Behav. Res.* 31:571–77

Shafer G. 1996. *The Art of Causal Conjecture*. Cambridge, MA: MIT Press

Skinner BF. 1976. *About Behaviorism*. New York: Vintage Books

Sobel ME. 1994. Causal inference in latent variable models. See von Eye & Clogg 1994, pp. 3–35

Steiger JH. 1979. Factor indeterminancy in the 1930s and the 1970s: some interesting parallels. *Psychometrika* 44:157–68

Steiger JH. 1996a. Dispelling some myths about factor indeterminacy. *Multivariate Behav. Res.* 31:539–50

Steiger JH. 1996b. Coming full circle in the history of factor indeterminancy. *Multivariate Behav. Res.* 31:617–30

Sutcliffe JP. 1965. A probability model for errors of classification. I. General considerations. *Psychometrika* 30:73–96

Tucker LR. 1958. Determination of parameters of a functional relation by factor analysis. *Psychometrika* 23:19–23

von Eye A, Clogg CC, eds. 1994. *Latent Variable Analysis*. Thousand Oaks, CA:Sage

Weisberg H. 1980. *Applied Linear Regression*. New York: Wiley

Wiley DE. 1973. The identification problem for structural equation models with unmeasured variables. In *Structural Equation Models in the Social Sciences*, ed. AS Goldberger, OD Duncan, pp. 69–83. New York: Academic

Willett JB, Sayer AG. 1994. Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychol. Bull.* 116:363–81

Annual Review of Psychology
Volume 53, 2002

CONTENTS

Analysis of Latent Variables

Indexes

Errata

Online log of corrections to the Annual Review of Psychology corrections:
Steven Regeser López and Peter J. Guarnaccia
Cultural Psychopathology: Uncovering the Social World of Mental Illness
Annu. Rev. Psychol. 2000, Vol. 51: 571–598.
http://psych.annualreviews.org/errata.shtml