On the Nature and Direction of Relationships Between Constructs and Measures

Jeffrey R. Edwards
University of North Carolina at Chapel Hill and
University of Michigan Business School

Richard P. Bagozzi
Rice University and University of Michigan
Business School

Theory development typically focuses on relationships among theoretical constructs, placing little emphasis on relationships between constructs and measures. In most cases, constructs are treated as causes of their measures. However, this causal flow is sometimes reversed, such that measures are viewed as causes of constructs. Procedures have been developed to identify and estimate models that specify constructs as causes or effects of measures. However, these procedures provide little guidance for determining a priori whether constructs should be specified as causes or effects of their measures. Moreover, these procedures address few of the possible causal structures by which constructs and measures may be related. This article develops principles for specifying the direction and structure of relationships between constructs and measures. These principles are illustrated using examples from psychological, sociological, and organizational research.

A theory can be divided into two parts: one that specifies relationships between theoretical constructs and another that describes relationships between constructs and measures (Bagozzi & Phillips, 1982; Costner, 1969). Presentations of theory often place great emphasis on explaining causal relationships among constructs but devote little attention to the nature and direction of relationships between constructs and

measures. These relationships are of paramount importance because they constitute an auxiliary theory that bridges the gap between abstract theoretical constructs and measurable empirical phenomena (Costner, 1969). Without this auxiliary theory, the mapping of theoretical constructs onto empirical phenomena is ambiguous, and theories cannot be meaningfully tested (Blalock, 1971).

The nature and direction of relationships between constructs and measures have been discussed in the literature on construct validity and structural equation modeling (Blalock, 1971; Bollen, 1989; DeVellis, 1991). In this literature, constructs are usually viewed as causes of measures, meaning that variation in a construct leads to variation in its measures (Bollen, 1989). Such measures are termed *reflective* because they represent reflections, or manifestations, of a construct (Fornell & Bookstein, 1982). Reflective measurement underlies classical test theory (Lord & Novick, 1968), reliability estimation (Nunnally, 1978), and factor analysis (Harman, 1976; Kim & Mueller, 1978), each of which treats a measure as a function of a latent variable (i.e., construct) plus error.

In some situations, measures are viewed as causes of constructs (Bagozzi & Fornell, 1982; Blalock, 1971; Bollen & Lennox, 1991; Heise, 1972; MacCallum & Browne, 1993). Such measures are termed *formative*, meaning the construct is formed or induced by

its measures (Fornell & Bookstein, 1982).[1] Formative measures are commonly used for constructs conceived as composites of specific component variables, as when socioeconomic status (SES) is defined in terms of occupation, education, and income (Hauser & Goldberger, 1971; Marsden, 1982).

Discussions of reflective and formative measures have focused primarily on identification and estimation issues (Blalock, 1971; Bollen & Lennox, 1991; MacCallum & Browne, 1993). These issues are critically important because they must be addressed before models with reflective and formative measures can be empirically tested. However, little attention has been devoted to the conditions in which measures should be specified as reflective or formative in the first place. Rough guidelines can be inferred from examples used to illustrate reflective and formative measures (Bollen & Lennox, 1991; MacCallum & Browne, 1993), but these examples are open to different interpretations and may not apply to measures used in a particular study. Covariances among measures can help differentiate formative from reflective measures, because these covariances follow a predictable pattern for reflective measures but are indeterminate for formative measures (Bollen & Lennox, 1991; Bollen & Ting, 1993). However, this approach permits only tentative conclusions, because imperfect reflective measures may yield covariances that deviate from the pattern expected for such measures, and formative measures may exhibit covariances that happen to follow the pattern expected for reflective measures. Moreover, a particular pattern of covariances may be consistent with numerous causal structures (Bollen & Ting, 1993; Duncan, 1975; Lee & Hershberger, 1990; MacCallum, Wegener, Uchino, & Fabrigar, 1993). Hence, researchers currently have few conceptual or empirical criteria for determining whether measures should be specified as reflective or formative. This shortcoming seriously hampers tests of substantive theories because if the causal structure relating constructs to measures is specified incorrectly, relationships among constructs cannot be meaningfully tested (Blalock, 1971; MacCallum & Browne, 1993).

This article provides general principles for specifying the nature and direction of the relationship between constructs and measures. These principles are derived from logical and philosophical arguments regarding the meaning of constructs and measures and the nature and form of their relationship (Blalock, 1971; Costner, 1969). These principles provide a priori criteria by which researchers can specify the causal direction and structure of the relationship between constructs and measures. By applying these principles, researchers may work toward developing auxiliary theories linking constructs to measures with the same precision and rigor used to specify relationships between theoretical constructs.

We begin by defining constructs and measures and addressing key philosophical issues underlying these definitions. Next, we draw from principles of causality to develop criteria for determining the direction of the relationship between a construct and a measure. We then derive hypothetical models that delinate basic causal structures by which a construct and a measure might relate. Cumulatively, this discussion yields guidelines for specifying the relationship between a construct and a measure in terms of (a) direction (i.e., whether a construct causes or is caused by its measures) and (b) structure (i.e., whether the relationship is direct, indirect, spurious, or unanalyzed). We then apply these guidelines to constructs and measures from the psychological, sociological, and organizational literatures. We conclude with recommendations for theory development, with particular emphasis on the integration of auxiliary theories relating constructs to measures into substantive theories of relationships among constructs.

Defining Constructs and Measures

Before we discuss relationships between constructs and measures, we must first define these terms. We define a *measure* as an observed score gathered through self-report, interview, observation, or some other means (DeVellis, 1991; Lord & Novick, 1968; Messick, 1995). Put simply, a measure is a quantified record, or datum, taken as an empirical analog to a construct. Note that a measure refers not to an instrument used to gather data or to the act of collecting data, but to the score generated by these procedures. A *construct* is a conceptual term used to describe a phe-

---

[1] Blalock (1964), who is often credited with distinguishing measures as effects versus causes of constructs, labeled reflective measures *effect indicators* and formative measures *cause indicators*. We prefer the terms *reflective* and *formative* because these terms do not imply that the relationship between a construct and a measure is necessarily causal. As we later show, relationship between some constructs and measures should be viewed as definitional rather than causal.

nomenon of theoretical interest (Cronbach & Meehl, 1955; Nunnally, 1978; Schwab, 1980). Several points regarding this definition should be elaborated. First, although constructs are terms researchers literally construct, or put together (Nunnally, 1978), we intend that constructs refer to phenomena that are real and exist apart from the awareness and interpretation of the researcher and the persons under study (Cook & Campbell, 1979; Loevinger, 1957; Messick, 1981). For instance, psychologists develop cognitive, attitudinal, and emotional constructs to describe real phenomena that are experienced by people. Likewise, sociologists formulate constructs such as social stratification, social mobility, and SES to represent real phenomena in social collectives. Second, although constructs refer to real phenomena, constructs themselves are not real in an objective sense (Nunnally, 1978). Rather, they are elements of scientific discourse that serve as verbal surrogates for phenomena of interest. Thus, when we speak of the relationship between a measure and a construct, we refer to the relationship between a measure and the phenomenon named by the construct. Third, the phenomena that constructs describe can be unobservable (e.g., attitudes) or observable (e.g., task performance). In either case, the construct itself is an abstract term that describes the phenomenon. Hence, subjective states are described as *attitudes* and behavior is described as *task performance* when researchers derive labels for these phenomena. Finally, constructs differ in how well they describe and assign meaning to phenomena of theoretical interest. Some constructs may demonstrate ongoing usefulness, whereas others initially considered useful may be modified or abandoned as knowledge accumulates. These advances may occur even when the phenomenon of interest remains unchanged. In sum, our definition of a construct represents a critical realist perspective, in that we view constructs as attempts to describe real phenomena, but we recognize that these phenomena cannot be known directly or with complete accuracy because of measurement error and the imperfect epistemological lens that a construct provides (Cook & Campbell, 1979; Delanty, 1997; Loevinger, 1957; Messick, 1981; Zuriff, 1998).

Causal Direction of the Relationship Between Constructs and Measures

Discussions of the relationship between constructs and measures often characterize this relationship as causal. Causal language pervades discussions of reflective measures. For example, Lord and Novick (1968) described measurement error as that part of an observed variable that is not "determined by" a construct (p. 531), and Nunnally (1978) characterized measurement error as the variance in a measure that is not "explained by" the true score (p. 201). Long (1983) was more explicit, stating that the arrow leading from a construct to its measure represents the "causal effect" of the construct on the measure (p. 11). Likewise, DeVellis (1991) asserted that a latent variable (i.e., a construct) is "a *cause* of the item score" (p. 13). Discussions of formative measures often use causal language as well. For example, Blalock (1971) described formative measures as indicating that a latent variable is measured using "one or more of its causes" (p. 336). Likewise, MacCallum and Browne (1993, p. 533) and Bedeian, Day, and Kelloway (1997, p. 788) stated that formative measures "may be viewed as causing" latent variables. Bollen and Lennox (1991) were more cautious, stating that they "do not attribute any special significance to the term *cause* other than the fact that the indicators determine the latent variable" (p. 306), although it is unclear how "determine" differs from the concept of cause.

Although discussions of formative and reflective measures often use causal language, they rarely incorporate principles of causality from the philosophy of science (e.g., Hume, 1946; Mill, 1886; Popper, 1959; Suppes, 1970). These principles are an ongoing source of debate, although the current literature suggests some consensus on four conditions for establishing causality in the social, behavioral, and management sciences (Asher, 1983; Bagozzi, 1980; Bollen, 1989; Cook & Campbell, 1979; Heise, 1975; James, Mulaik, & Brett, 1982). First, causality requires that the cause and the effect are distinct entities. When two variables are not distinct, their relationship is tautological and therefore should not be viewed as causal. Second, causality requires association, meaning that the cause and the effect covary. Association is usually viewed as probabilistic rather than definitional, meaning the cause increases the likelihood of the effect but does not guarantee that the effect will occur (Cook & Campbell, 1979; James et al., 1982; Suppes, 1970). Third, causality requires temporal precedence, such that the cause occurs before the effect. Although causality may be nearly instantaneous at the micromedial level, the cause must precede the effect by some minimal time interval (Cook & Campbell, 1979). Finally, causality requires

the elimination of rival explanations for the presumed relationship between the cause and the effect (Cook & Campbell, 1979). Of the four conditions for causality, this last condition is often the most difficult to satisfy in practice, particularly in nonexperimental research. We now apply the foregoing conditions for causality to the relationship between constructs and measures.

Distinct Entities

The first condition for causality stipulates that a construct and a measure must be distinct. This condition is satisfied by the definitions of construct and measure provided earlier, in which a construct refers to a phenomenon of theoretical interest and a measure is an observed score. The distinction between a construct and a measure is evident for attitudes, cognitions, and other mental states or events, for which the construct refers to a phenomenon within the mind of a person and the measure is a recorded trace taken as evidence of the construct (e.g., a number circled by a respondent on a survey). Constructs that refer to behavior are also distinct from measures of behavior, because exhibiting a behavior is not the same as the score associated with that behavior. For example, the act of being absent from work is distinct from a score on the item "How many days of work have you missed in the past year?" (Johns, 1994). The distinction between a construct and a measure breaks down under operationalism, in which a construct is defined in terms of its measures; for example, *intelligence* was once defined as that which the Stanford–Binet measures (Campbell, 1960).

Association

The second condition for causality is association, meaning the construct and the measure must covary. In philosophical discussions of causality, association usually refers to an empirical relationship between two objects. For a construct and a measure, however, the researcher has direct access to the measure but not the construct (Nunnally, 1978). Hence, association between a construct and a measure must be inferred by using methods that do not rely on direct observation.

One such method involves the use of covariances among multiple measures of a construct to infer the relationships between the focal construct and its measures. For example, if a construct has three or more reflective measures, the covariances among the measures may be used to estimate the association (i.e., factor loading) between each measure and the con-

struct (Bollen, 1989). Alternatively, if a construct has one or more formative measures, the relationships between the measures and the construct can be estimated provided the construct is specified as a direct or an indirect cause of at least two reflective measures (Bollen & Davis, 1994). Although these procedures provide empirical evidence of association, this evidence may be consistent with other models with different causal orderings that yield different estimates of the relationships between constructs and measures (Duncan, 1975; Lee & Hershberger, 1990; MacCallum et al., 1993). Therefore, empirical evidence of association provides necessary but not sufficient support for a hypothesized causal relationship between a construct and a measure.[2]

A second method for evaluating the association between a construct and a measure entails the use of "mental experiments" (Bollen, 1989), in which a researcher imagines a change in the construct and then considers whether a change in the measure is likely. For example, a researcher might consider whether variation in a job involvement construct would relate to different scores for items such as "The most important things that happen in life involve work" (Kanungo, 1982). Mental experiments entail the use of disciplined imagination (Weick, 1989) to deduce whether an association between a construct and a measure is plausible. Ultimately, mental experiments rely on speculation and "appeals to reason" that a construct and a measure covary (cf. Nunnally, 1978, p. 93). Therefore, mental experiments should be used in conjunction with empirical procedures, acknowledging that neither approach provides definitive evidence of association between a construct and a measure.

Ideally, the association between a construct and its measures should remain stable regardless of the larger causal model in which the construct and measures are embedded. If the association varies, then no unique meaning can be assigned to the construct (Burt, 1976), and claims of association between the construct and measures are tenuous. The association between a construct and its measures is generally stable for reflec-

---

[2] We refer to association *net of other influences* as a necessary condition for causality (Bollen, 1989, p. 57). A bivariate association between a construct and a measure is not a necessary condition for causality because their association may entail two effects of opposite sign that yield a null total effect.

tive measures that serve as alternative (i.e., substitute) indicators of a construct, provided these measures correlate more highly with one another than with measures of other constructs. In contrast, associations of formative measures with their construct are determined primarily by the relationships between these measures and measures of constructs that are dependent on the construct of interest. This point was made by Heise (1972), who noted that a construct measured formatively is not just a composite of its measures; rather, "it is the composite that best predicts the dependent variable in the analysis. . . . Thus, the meaning of the latent construct is as much a function of the dependent variable as it is a function of its indicators" (p. 160). Unstable associations between formative measures and constructs not only create difficulties for establishing causality but also obscure the meaning of these constructs because their interpretation depends on the dependent variables included in a given model.

Temporal Precedence

The third condition for causality is temporal precedence, or whether change in the construct precedes, accompanies, or follows change in the measure. Because researchers have direct access to measures but not constructs, temporal precedence cannot be assessed directly. However, it may be established by using experiments that control the timing and sequence of a construct and its measures. To create a sequence in which the construct precedes the measure, researchers may apply experimental manipulations known to influence the construct and collect scores after the manipulation. For example, studies of perceived control have shown that giving participants a switch that terminates electric shock increases perceived control (Averill, 1973). If this manipulation is followed by a self-report measure of control, one may conclude that variance in the perceived-control construct preceded variance in scores on the measure of control. To create a sequence in which the measure precedes the construct, scores may be manipulated by providing bogus feedback to subjects, under the assumption that variation in these scores will create variation in a latent attitude or belief construct (Binning, Zaba, & Whattam, 1986; Phillips & Lord, 1982). However, to verify that the construct has indeed changed, a valid reflective measure of the construct must follow. Both of these experimental procedures entail a stimulus → organism → response

sequence in which one link in the sequence must be taken as given (i.e., the stimulus → organism link must be assumed to test the effect of a construct on a measure, whereas the organism → response link must be assumed to test the effect of a measure on a construct).

If conventional experimentation is not possible, researchers may use mental experiments (Bollen, 1989) to evaluate various temporal orderings between a construct and a measure. For example, constructs representing well-formed attitudes presumably exist before being measured by a researcher. Such constructs should therefore precede their measures, indicating a temporal sequence consistent with reflective measurement. Some attitudes may be formed or changed in response to the act of measurement (Feldman & Lynch, 1988), which may invite the conclusion that a measure precedes its construct. However, studies of these effects entail a temporal sequence in which presentation of an item is followed by an attitude, which in turn is followed by an emitted score. Although the presentation of the item precedes the formation or change of the construct, the construct precedes its associated score. This reasoning draws from our definition of a measure as the score itself, not the device used to gather the score.

Some researchers have used mental experiments to argue that formative measures precede their constructs. For example, Heise (1972) argued that the construct SES is caused by measures of education, income, and occupational prestige on the basis of the premise that changes in these socioeconomic variables lead to changes in SES, but not the reverse. Although this logic is appealing, it relies on a subtle form of operationalism in which measures of education, income, and occupational prestige are equated with the socioeconomic phenomena they are intended to represent. Thus, a score denoting years of schooling is treated as education itself, not as an indicator of education containing errors due to imperfect recall, coding mistakes, and so forth. If we assume that scores on education, income, and occupational prestige contain measurement error and occur after the phenomena they represent (e.g., reporting one's years of schooling logically occurs after the last year of schooling has been completed), then these scores may be viewed as reflective measures of socioeconomic constructs that in turn cause SES. Following this logic, it is these socioeconomic constructs, not their measures, that cause SES (Blalock, 1971).

Eliminating Rival Causal Explanations

The fourth condition for causality is the elimination of rival explanations for the presumed causal relationship between a construct and a measure. This condition is perhaps the most difficult to satisfy because of the myriad third variables that may induce a relationship between a presumed cause and effect (Cook & Campbell, 1979). Moreover, rival causal explanations must be examined with regard to specific constructs and measures (Hauser, 1972). Hence, ruling out rival causal explanations is a daunting task that cannot be reduced to universal prescriptions. Thus, our objective here is to suggest general procedures for identifying rival explanations for a presumed causal relationship between a construct and a measure and to identify broad categories of rival explanations that may apply in a variety of contexts.

One approach to identifying rival causal explanations for constructs and measures is to consider threats to validity relevant to quasi-experimental research (Cook & Campbell, 1979). One threat to internal validity is history, in which the relationship between the presumed cause and effect is due to some intervening event that is not part of the experimental treatment. Applied to constructs and measures, history entails some event that occurs between the construct and the measure that may explain their relationship. This threat to validity was illustrated by Brief, Butcher, and Roberson (1995), who examined job satisfaction in hospital employees. In general, one might assume that variation in the job satisfaction construct would generate variation in job satisfaction scores, indicating reflective measurement. However, Brief et al. showed that positive mood-inducing events at the time of measurement (i.e., giving respondents small gifts, such as cookies) influenced job satisfaction scores. These events are a historical threat to validity because they occurred between the construct and the measure (assuming employees had some notion of their job satisfaction before the study) and influenced the measure irrespective of the job satisfaction construct.[3]

Another threat to validity that is relevant to constructs and measures is instrumentation, in which the relationship between a presumed cause and effect is due in part to variation in the instrument or method used to collect data (Cook & Campbell, 1979). For example, job performance scores from a set of raters may vary not only because job performance varies across ratees but also because raters use different rating procedures (Coovert, Craiger, & Teachout, 1997). Thus, rater variance provides a rival explanation for

the presumed effect of job performance on ratings. This effect is a special case of method variance as a nuisance factor in observed scores (Bagozzi & Yi, 1990; Williams & Brown, 1994). Additional threats to validity (Cook & Campbell, 1979) may be considered to assess other rival explanations for relationships between constructs and measures.

A related approach to identifying rival causal explanations for constructs and measures is to apply mental experiments, as discussed earlier. For example, overall job satisfaction is often measured using items that describe satisfaction with specific job facets (e.g., Warr, Cook, & Wall, 1979; Weiss, Dawis, England, & Lofquist, 1967). Scores on these items are then taken as direct reflective measures of overall job satisfaction, as implied by the use of statistical procedures (e.g., reliability estimation) consistent with reflective measurement (Bollen & Lennox, 1991; Nunnally, 1978). However, the presumed effect of the overall job satisfaction construct on facet satisfaction measures may be spurious, attributable to facet satisfaction constructs. This rival explanation is based on two plausible assertions: (a) Facet satisfaction measures are directly influenced by facet satisfaction constructs, not overall job satisfaction (Ironson, Smith, Brannick, Gibson, & Paul, 1989), and (b) facet satisfaction constructs affect the overall job satisfaction construct, meaning that one's overall attitude toward a job arises from attitudes toward specific aspects of the job (Aldag & Brief, 1978; Ferratt, 1981). This logic can be applied to other general constructs operationalized using measures of specific aspects of those constructs.

Models of the Relationship Between Constructs and Measures

The preceding discussion outlined general principles for determining the causal direction between a construct and a measure. In this section, we build on those principles by developing formal models of the causal structure relating constructs to measures. His-

---

[3] A subtle difference between Brief et al.'s (1995) study and history as described by Cook and Campbell (1979) is that the positive mood-inducing events were not confounded with some other treatment intended to influence job satisfaction. Nonetheless, these events aptly demonstrate how conditions unrelated to the construct of interest can intercede the construct and its measure and lead to erroneous inferences regarding their relationship.

torically, models relating constructs to measures have focused on direct effects (Blalock, 1964; Bollen & Lennox, 1991; Costner, 1969; MacCallum & Browne, 1993). We extend these models by applying rules for decomposing correlations in path analysis and structural equation modeling (Alwin & Hauser, 1975; Fox, 1980). These rules stipulate that the correlation between two variables can be decomposed into four components: (a) a *direct effect*, in which one variable directly affects another; (b) an *indirect effect*, in which the effect of one variable on another is mediated by (i.e., transmitted through) one or more other variables; (c) a *spurious component* that is due to common or correlated causes; and (d) an *unanalyzed component* resulting from associations among predetermined (i.e., exogenous) variables. Direct and indirect effects can operate in either direction, depending on which variable is specified as the cause of the other. Applying these distinctions to the relationship between a construct and a measure yields six basic categories of models, shown later in Figures 1–6.

To simplify our presentation of these models, we made the following general assumptions and restrictions: (a) Relationships in the models are linear, (b) constructs and measures are expressed as deviations from their means, (c) measurement errors and disturbance terms are random and are therefore uncorrelated with one another and with predictors in the equation that contains the measurement error or disturbance term, (d) each construct has three measures, (e) models with indirect effects contain a single stage of mediation, (f) causal flow between the construct and the measure is unidirectional (i.e., recursive), and (g) measures cannot directly cause other measures (i.e., scores are inert quantitative symbols and therefore cannot directly influence one other).[4] We later suggest how the simplified models considered here can be combined to represent more complex causal mechanisms linking constructs to measures. The notation for all models follows conventions from the structural equation modeling literature (e.g., Jöreskog & Sörbom, 1996).

Our discussion of these models does not address issues of identification, and some of the models we discuss are not identified as presented. However, our goal here is to derive models that are logically possible, independent of whether they are identified. This goal reflects the premise that correct model specification should take priority over identification. Unidentified models can become identified by constraining parameters or by adding reflective measures of the
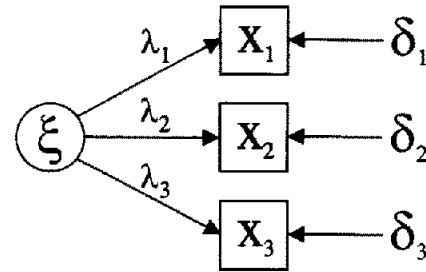

Figure 1. Direct reflective model.

focal construct or its effects (Bollen & Davis, 1994; MacCallum & Browne, 1993). However, these modifications should not alter the specified causal flow between the construct and its measures. If this causal flow is altered, the result is a model that is empirically testable but incorrectly represents the causal relationships of interest. The usefulness of testing such a model is dubious at best.

Direct Reflective Model

The direct reflective model specifies direct effects from a construct to its measures. This model is depicted in Figure 1, in which each of the x_i measures is influenced by the construct ξ and the random measurement error δ_i. Hence, variance in each measure is explained by a construct common to all measures and error unique to each measure, and covariation among the measures is attributed to their common cause, ξ. The direct reflective model is expressed by the following equation:

$$x_i = \lambda_i \xi + \delta_i. \tag{1}$$

In Equation 1, the λ_i are factor loadings depicting the magnitude of the effect of ξ on the x_i, and δ_i represents random measurement error (i ranges from 1 to 3 for the model in Figure 1). Equation 1 has its roots in classical test theory (Lord & Novick, 1968; Nunnally,

---

[4] Conceivably, models could be specified in which measures were the causes of other measures. For example, in a model specifying an indirect effect of a construct on a measure, the mediating variable could, in principle, be either a construct or a measure. However, the relationship of one measure with another is generally limited to deterministic functional relationships, as when one score is mathematically transformed into another score (e.g., logarithm, square root) or multiple scores are summed to create scales or item parcels. These relationships are not causal; rather, they simply entail mathematical operations on inert data.

1978) and underlies reliability estimation, common factor analysis, and confirmatory factor analysis (Bollen, 1989; Harman, 1976; Kim & Mueller, 1978; Long, 1983; Nunnally, 1978).

Direct Formative Model

The direct formative model specifies measures as correlated causes of a construct. This model is depicted in Figure 2, which shows effects from the x_i measures to the construct η. The disturbance term ζ represents that part of the construct η that is not explained by the x_i measures and thus may be interpreted as measurement error. In contrast, the x_i are conceived as error-free causes of η (MacCallum & Browne, 1993). Because causality runs from the measures to the construct, the construct η is not presumed to explain the variances of the x_i or the covariances among the x_j. The direct formative model is represented by the following equation:

$$\eta = \sum_i \gamma_i x_i + \zeta, \qquad (2)$$

where the γ_i are structural parameters depicting the magnitudes of the effects of the x_i on η, and Σ represents the summation of the $\gamma_i x_i$ products (i ranges from 1 to 3 for the model in Figure 2).

Because the direct formative model may be unfamiliar to the reader, we describe several of its common uses. One use of the model is to create an "induced" latent variable that represents an aggregation of observed variables (Heise, 1972). An oft-cited example is SES, which is typically viewed as a function of background variables such as income, education, and occupational prestige (Bollen & Lennox, 1991; Heise, 1972). Other induced latent variables include stress as measured by the experience of major life events (Cohen, Cohen, Teresi, Marchi, & Velez, 1990); social support as measured by number of supportive incidents (MacCallum & Browne, 1993); and
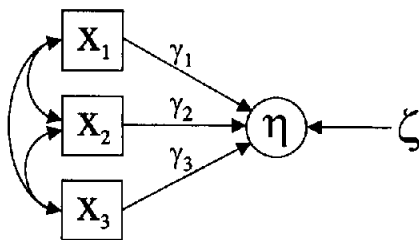

Figure 2. Direct formative model (for simplicity, $\phi..$ labels on covariances among exogenous variables are omitted).

exposure to discrimination as measured by race, age, sex, and disabilities (Bollen & Lennox, 1991).

In some cases, the equation relating an induced latent variable to its measures excludes the disturbance term ζ, such that the latent variable is a weighted linear function of its measures. This use of the direct formative model underlies principal-components analysis (Kim & Mueller, 1978), canonical correlation analysis (Thompson, 1984), and partial least squares (Garthwaite, 1994; Wold, 1982), each of which uses observed measures to create weighted linear composites that serve as conceptual variables in subsequent analyses. By omitting the disturbance term, measurement error in these composites is effectively ignored.

A second use of the direct formative model is to create block variables that summarize the effects of several variables (Coleman, 1976; Heise, 1972; Igra, 1979; Marsden, 1982). A block variable yields a single summary estimate of the effects of variables in the block on some outcome. Variables in the block often represent conceptually distinct causes of the outcome. For example, in a model predicting political liberalism, Heise (1972) introduced a block variable labeled *family socialization*, which itself was a function of mother's liberalism, father's liberalism, and other unspecified variables captured by the disturbance term ζ. The overall effect of family socialization on political liberalism was summarized with a "sheaf coefficient," which was calculated from coefficients obtained by regressing political liberalism on mother's and father's liberalism.

Third, the direct formative model has been used to depict the effects of an experimental manipulation on a latent variable (Alwin & Tessler, 1974; Bagozzi, 1977; Blalock, 1971; Costner, 1971). For example, Costner (1971) described an experiment in which a construct signifying fatigue was manipulated by withholding sleep from participants. Here, degree of sleep deprivation was taken as a measure of fatigue. If an experiment includes measures designated as manipulation checks, these measures may be viewed as reflective indicators of the manipulated construct (Bagozzi, 1977; Costner, 1971).

Indirect Reflective Model

The preceding models specified relationships between a construct and its measures as direct. We now consider two models that capture indirect effects. One is the indirect reflective model (see Figure 3). This model shows that the effects of the construct ξ on its
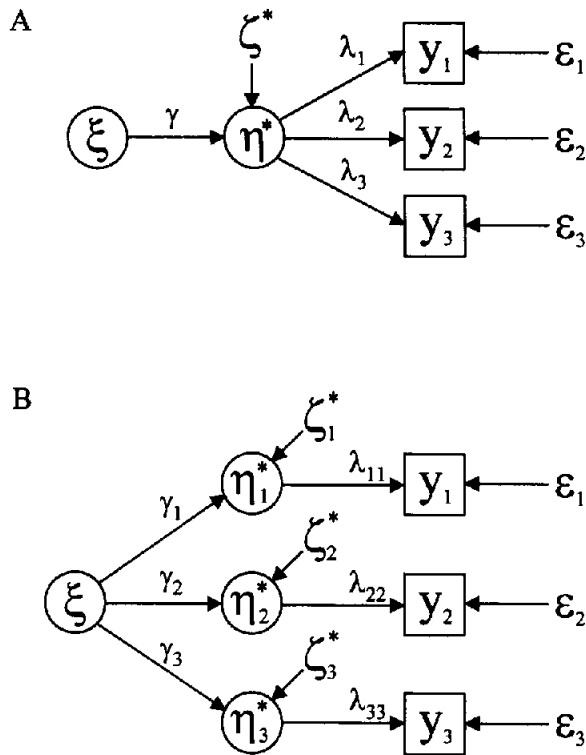
A

B

Figure 3. Indirect reflective model. A: Single mediating construct. B: Multiple mediating constructs.

y_i measures are mediated by one or more latent variables, η_j^*. We use an asterisk to indicate that the η_j^* are not the construct intended to be measured by the y_i, which instead is labeled ξ. The indirect reflective model applies when measures are interpreted as indicators of a construct but actually represent one or more of its effects. For example, job satisfaction has been measured with items describing turnover intent (e.g., "I frequently think of quitting my job"; Hackman & Oldham, 1980). However, theories of job satisfaction and turnover indicate that turnover intent is a consequence of job dissatisfaction, not the affective experience of dissatisfaction itself (Locke, 1976; Mobley, 1982).

Figure 3 shows two versions of this model. The first incorporates a single mediating construct, such that ξ influences a single η^*, which in turn affects multiple y_i.[5] The second version contains multiple mediating constructs (i.e., η_j^*), each of which affects a single y_i. Note that, for both versions of the model, the η^* or η_j^* are also affected by disturbance term ζ^* or ζ_j^*, respectively, and the y_i are influenced by measurement errors designated as ε_i. A general set of

equations that capture both versions of the indirect reflective model can be written as follows:

$$\eta_j^* = \gamma_j \xi + \zeta_j^* \tag{3}$$

and

$$y_i = \lambda_{ij}\eta_j^* + \varepsilon_i. \tag{4}$$

In Equation 3, γ_j indicates the effect of ξ on the η_j^*, and in Equation 4, λ_{ij} captures the effects of the η_j^* on the y_i. For both models shown in Figure 3, i ranges from 1 to 3; for the model in Figure 3A, $j = 1$ (and is therefore disregarded; see footnote 5), whereas for the model in Figure 3B, j ranges from 1 to 3. Note that Equation 4 accommodates cross-loadings, meaning that each y_i may load on each η_j^*. However, the model in Figure 3B depicts a special case of Equation 4 in which each y_i loads on a single η_j^*, meaning that each y_i represents its own mediating constructs. Equations 3 and 4 correspond to a second-order factor model in which the y_i load on the first-order factors η_j^*, which in turn load on the second-order factor ξ (Rindskopf & Rose, 1988). However, unlike most applications of second-order factor analysis, our emphasis here is on cases in which y_i are mistaken as indicators of the construct ξ.

An important insight regarding the relationship between ξ and the y_i can be obtained by substituting the expression for η_j^* from Equation 3 into Equation 4:

$$y_i = \lambda_{ij}(\gamma_j \xi + \zeta_j^*) + \varepsilon_i$$
$$= \lambda_{ij}\gamma_j \xi + \lambda_{ij}\zeta_j^* + \varepsilon_i. \tag{5}$$

Equation 5 shows that the relationships between the y_i and the construct they are intended to measure (i.e., ξ) are represented by the products $\lambda_{ij}\gamma_j$. Therefore, any estimate of the relationships between ξ and the y_i nec-

---

[5] For simplicity, when a model contains a single variable of a particular type, we omit the subscript for that variable. For example, because the model in Figure 3A contains a single η^*, we omit the j subscript in our discussion of η^* for this model, and we do not include a subscript for this variable in Figure 3A (the implied subscript of 1 is superfluous, given that there is no need to index this variable). By the same token, parameters linked to a single variable of a particular type do not include a subscript in reference to that variable. Thus, the λ parameters in Figure 3A contain subscripts for the y_i but not for the single η^*. When a parameter links two variables that do not require subscripts, the parameter likewise has no subscripts (as exemplified by the γ linking ξ and η^* in Figure 3A).

essarily confounds the relationships between ξ and the η_j^* with the relationships between the η_j^* and the y_i. Equation 5 also shows that variance in the y_i is attributable not only to the construct ξ and the measurement errors ε_j but also to the disturbance terms ζ_j^*. The effects of the ζ_j^* on the y_i represent additional sources of measurement error in the y_i, given that the ζ_j^* represent that portion of the variance in the η_j^* not explained by ξ and hence signify error.[6]

Indirect Formative Model

A second model that incorporates indirect effects between constructs and measures is the indirect formative model, shown in Figure 4. Here, the x_i have a direct impact on one or more η_j^*, which in turn influence the construct of interest, η. This model applies when formative measures assigned to a construct actually represent one of its causes. For example, Mac-Callum and Browne (1993, p. 538) showed that a model using supportive incidents as direct formative
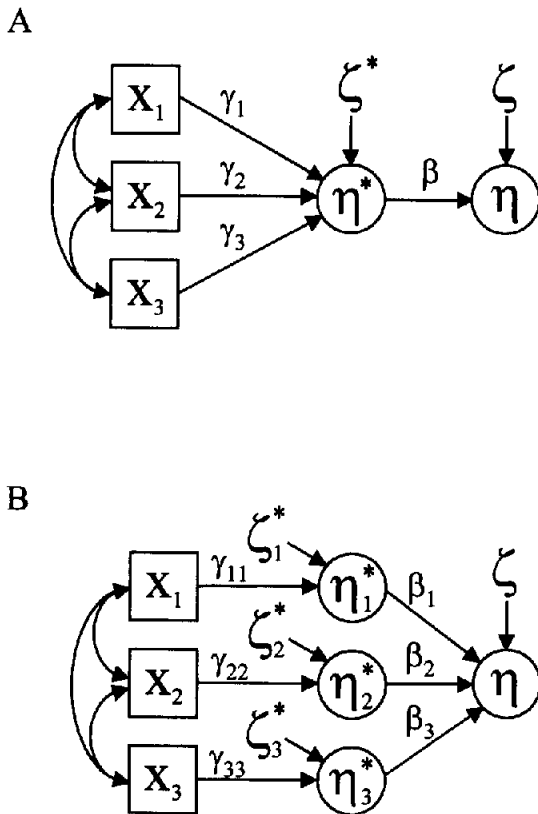
A

B

Figure 4. Indirect formative model (for simplicity, ϕ.. labels on covariances among exogenous variables are omitted). A: Single mediating construct. B: Multiple mediating constructs.

measures of depression could be recast as an indirect formative model in which supportive incidents created an induced social support variable, which in turn influenced depression. We consider two versions of the indirect formative model: one in which the x_i influence a single η^*, which in turn affects η, and another in which the x_i influence multiple η_j^* (in this example, three), each of which affects the η. Note that the η^*, η_j^*, and η are also influenced by corresponding disturbance terms ζ^*, ζ_j^*, and ζ. Both versions of this model can be represented by the following two equations:

$$\eta_j^* = \sum_i \gamma_{ji} x_i + \zeta_j^* \tag{6}$$

$$\eta = \sum_j \beta_j \eta_j^* + \zeta. \tag{7}$$

In Equation 6, the γ_{ji} depict the magnitudes of the effects of the x_i on the η_j^*, and Σ is the summation of the $\gamma_{ji} x_i$ products over i ($i = 1$ to 3 for both examples shown in Figure 4). In Equation 7, the β_j indicate the magnitudes of the effects of the η_j^* on η, and Σ is the summation of the $\beta_j \eta_j^*$ products over j ($j = 1$ for the model in Figure 4A and is therefore omitted; $j = 1$ to 3 for the model in Figure 4B). The relationships between the construct of interest, η, and its measures x_i may be further clarified by substituting the expression for η_j^* in Equation 6 into Equation 7:

$$\eta = \sum_j \beta_j \left(\sum_i \gamma_{ji} x_i + \zeta_j^* \right) + \zeta$$

$$= \sum_j \sum_i \beta_j \gamma_{ji} x_i + \sum_j \beta_j \zeta_j^* + \zeta. \tag{8}$$

Equation 8 reveals that the relationship between the x_i and the intended construct, η, is represented by the summation of the $\beta_j \gamma_{ji}$ across all η_j^*. Thus, an estimate of the effect of x_i on η confounds the effects of x_i on the η_j^* with the effects of the η_j^* on η. In addition, variance in η is attributable not only to its measures x_i and the disturbance ζ but also to the disturbances ζ_j^*. Thus, measurement error in η is represented by ζ as well as the products $\beta_j \zeta_j^*$ associated with each of the η_j^*.

---

[6] The interpretation of the ζ_j^* as error is based on the assumption that the y_i are intended to measure ξ rather than the η_j^*. If y_i were intended to measure η_j^*, measurement error would be represented solely by the ε_j and the ζ_j^* would be interpreted simply as that part of the true score variance in the η_j^* not explained by ξ.
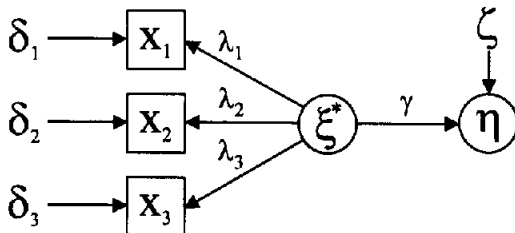
Spurious Model

The preceding models were based on the premise that, either directly or indirectly, the construct caused its measures or the measures caused the construct. An alternative model depicts the relationships between the construct and measures as spurious, due to the influence of one or more common causes (Blalock, 1971; Costner, 1969). Figure 5 displays two versions of this model. In the first version, covariances between the construct of interest, η, and its intended measures, x_i are due to a single common cause, ξ^*. The disturbance ζ indicates that the relationship between ξ^* and η is imperfect, such that some portion of the variance in η is due to forces other than ξ^*. In the second version, covariation between η and each x_i is due to a separate common cause, ξ_j^* (j ranges from 1 to 3). Equations corresponding to both versions of this model can be written as follows:

$$x_i = \lambda_{ij}\xi_j^* + \delta_i \tag{9}$$

$$\eta = \sum_j \gamma_j \xi_j^* + \zeta. \tag{10}$$

In Equation 9, λ_{ij} designates the effects of the ξ_j^* on the x_i, and in Equation 10, γ_j indicates the effect of ξ_j^*
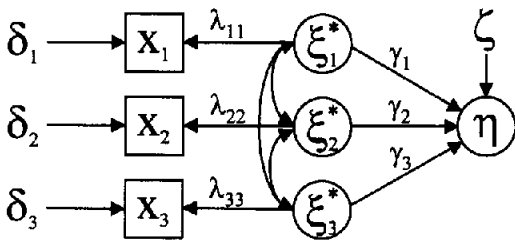
A

B

Figure 5. Spurious model (for simplicity, ϕ.. labels on covariances among exogenous variables are omitted). A: Single common cause. B: Multiple common causes.

on η. Assuming the δ_i and ζ are independent, a general expression for the covariance of the construct η with its ith measure x_i can be written as

$$\sigma(x_i, \eta) = \sum_j \sum_k \lambda_{ij}\phi_{jk}\gamma_k. \tag{11}$$

In Equation 11, i indicates the measure under consideration, j indexes the linkage between the measure and a particular ξ_j^*, k indexes the linkage between a particular ξ_k^* and η, and ϕ_{jk} refers to the variances and covariances of ξ_j^* and ξ_k^* (separate subscripts for ξ^* are required because some elements of the covariance between x_i and η involve different ξ^*). Equation 11 allows for cross-loadings, meaning that each of the x_i measures can be assigned to each of the ξ_j^* constructs. If each item is assigned to a single ξ_j^* construct, as in the model shown in Figure 5B, Equation 11 simplifies to

$$\sigma(x_i, \eta) = \lambda_{ij}\sum_k \phi_{jk}\gamma_k. \tag{12}$$

In Equation 12, i again indicates the measure in question, j represents the construct to which the measure is assigned, and k indexes each of the ξ_k^* constructs. For illustration, we apply this equation to the covariance between x_1 and η in Figure 5B, which yields the following:

$$\sigma(x_1, \eta) = \lambda_{11}\phi_{11}\gamma_1 + \lambda_{11}\phi_{12}\gamma_2 + \lambda_{11}\phi_{13}\gamma_3. \tag{13}$$

The first term in Equation 13 represents that part of the covariance between x_1 and η that is due to their common cause, ξ_1^*, whereas the latter two terms represent parts of the covariance between x_1 and η that is due to correlated causes (i.e., ξ_1^* and ξ_2^*, ξ_1^* and ξ_3^*, respectively). This expression shows that the relationship between x_1 and η confounds the relationship of x_1 with ξ_1^*, the relationship of each ξ_k^* with η, and the correlations among the ξ_k^*.

If a single common cause ξ^* is involved (as in Figure 5A), Equation 11 further simplifies to

$$\sigma(x_i, \eta) = \lambda_i\phi\gamma. \tag{14}$$

Hence, the covariance between x_1 and η in Figure 5A is simply $\lambda_1\phi\gamma$. This expression shows that the relationship between x_1 and η confounds the relationship of x_1 with ξ^* with the relationship between ξ^* and η.

A noteworthy special case of the spurious model with multiple common causes (Figure 5B) treats the x_i as error-free measures of the ξ_j^*, such that the λ_{ij} are unity and the variances of the δ_i are zero. In this case, the x_i and ξ_j^* are equivalent, and Equation 9 reduces to
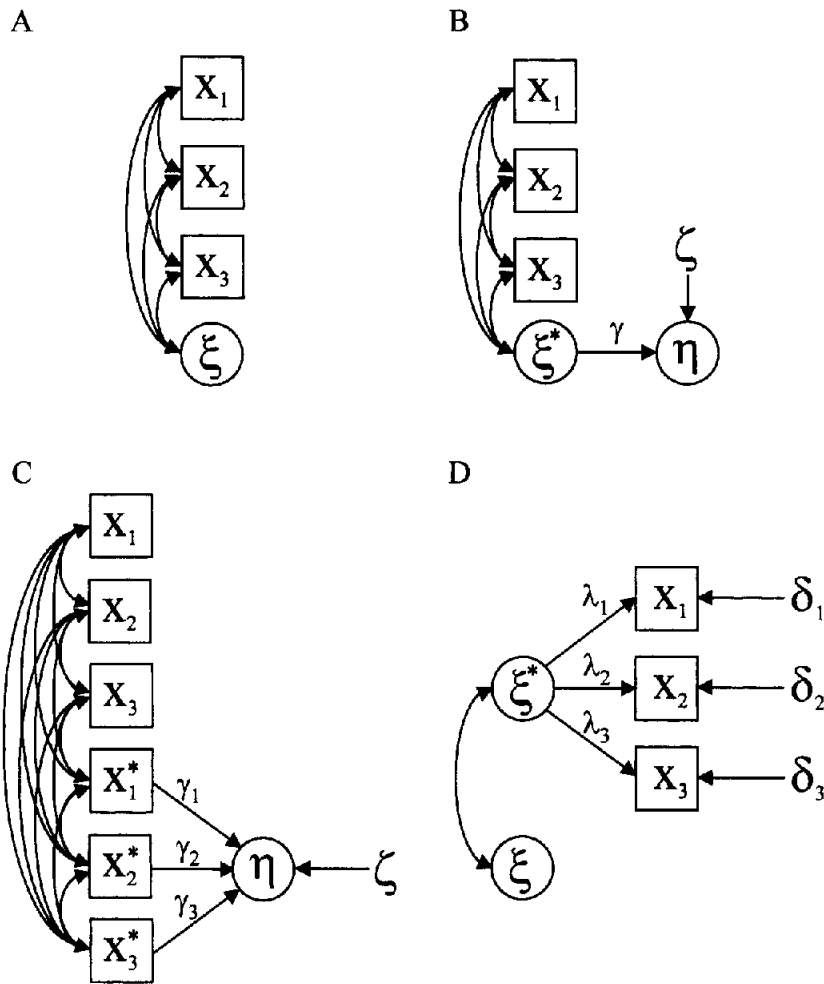
Figure 6. Unanalyzed model (for simplicity, ϕ.. labels on covariances among exogenous variables are omitted). A: Constructs and measures exogenous. B: Exogenous measures with correlated construct. C: Exogenous measures with correlated measures. D: Exogenous construct with correlated construct.

$x_i = \xi_j^*$. Substituting x_i for ξ_j^* in Equation 10 (and modifying the subscript on γ_j to correspond to the x_i) yields

$$\eta = \sum_i \gamma_i x_i + \zeta. \qquad (15)$$

Equation 15 is identical to the direct formative model as specified in Equation 2. Hence, the direct formative model may be considered a special case of the spurious indicator model in which the x_i are viewed as perfect (i.e., error-free) indicators of the ξ_j^*.

Unanalyzed Model

Finally, the relationships between a construct and its measures can be unanalyzed. This model applies when (a) the construct and measures are exogenous,

(b) the measures are exogenous and correlated with a construct that causes the focal construct, (c) the measures are exogenous and correlated with other measures that cause the construct, or (d) the focal construct is exogenous and correlated with another construct that causes the measures.[7] These distinctions yield four versions of the unanalyzed model, shown in Figure 6. For each model, we presume the researcher does not realize that the unanalyzed measures are indicators of the focal construct.

---

[7] For reasons stated earlier, we excluded models in which measures cause measures (for the unanalyzed case, this would entail a model in which a construct is correlated with a set of measures, which in turn cause other measures).

The model in Figure 6A specifies that the construct and its measures are exogenous, meaning their relationships are explained by forces outside the model. This model applies when the researcher is unwilling or unable to specify the relationship between a construct and its measures. Hence, the relationships between the construct and its measures are represented by the covariances $\phi_{i\xi}$ (here, the subscripts i and ξ identify the covariance between each x_i and ξ, thereby disregarding the covariances among the x_i). Because all variables are exogenous, the model does not attempt to explain variance in the x_i or ξ. Thus, the model cannot distinguish true score variance from measurement error in the x_i, nor can the model identify residual variance in ξ.

The model in Figure 6B indicates that the unanalyzed x_i measures are correlated with a construct, ξ^*, that causes the construct of interest, η. Here, the x_i are mistakenly interpreted as correlational measures of η, when in fact they represent correlational measures of ξ^*, which in turn causes η. The covariance between each measure and the intended construct can be expressed as

$$\sigma(x_i, \eta) = \gamma\phi_{i\xi*}, \qquad (16)$$

where γ signifies the effect of ξ^* on η, and $\phi_{i\xi*}$ represents the covariance between the ith measure and the construct ξ^*. Hence, the covariance between the x_i and η confounds the covariance between the x_i and ξ^* with the effect of ξ^* on η.

According to the model in Figure 6C, the x_i are viewed as correlational measures of η but in fact are correlated with a second set of measures, x_h^* (where h ranges from 1 to 3), which in turn are formative measures of η. The covariance between each x_i measure and the intended construct is represented as follows:

$$\sigma(x_i, \eta) = \sum_h \phi_{ih}\gamma_h. \qquad (17)$$

In Equation 17, x_i represents the measure of interest, and h indexes the covariance between x_i and each x_h^* and the effect of each x_h^* on η. This equation shows that the covariance between the x_i and η confounds the covariance between the x_i and each x_h^* and the effect of each x_h^* on η.

Finally, the model in Figure 6D indicates that the unanalyzed x_i are considered correlational measures of ξ but are actually reflective measures of ξ^*, which signifies a construct correlated with ξ. This model applies when a measure reflects a construct that is correlated with, but conceptually distinct from, the construct of interest. The covariance between each measure and the intended construct can be written as follows:

$$\sigma(x_i, \xi) = \phi_{\xi\xi*}\lambda_i, \qquad (18)$$

where $\phi_{\xi\xi*}$ is the covariance of ξ and ξ^*, and λ_i is the loading of x_i on ξ^*. Thus, the relationship of ξ with a given x_i confounds the covariance between ξ and ξ^* and the loading from ξ^* to x_i.

Summary and Extensions

The preceding models are summarized in Table 1, which classifies the models according to whether (a) the measure refers to the construct itself or to a cause or an effect of the construct and (b) the measure is reflective or formative with respect to the construct to which the measure is directly linked (i.e., a causal construct, the construct of interest, or an effect construct, moving left to right across Table 1). As a general rule, if a measure describes the inherent attributes of a construct, the relationship between the construct and the measure should be considered direct, whereas

Table 1
Classification of Models Relating Constructs to Measures

| Type of measure | Referent of measure | | |
	Cause of construct	Inherent attributes of construct	Effect of construct
Reflective	Spurious model	Direct reflective model	Indirect reflective model
Formative	Indirect formative model	Direct formative model	Unanalyzed model

Note. The columns indicate the referent of the measure, meaning whether the proximal construct (i.e., the construct to which the measure directly refers) is a cause of the construct of interest, the construct of interest itself, or an effect of the construct of interest. The rows indicate whether the measure is reflective or formative, meaning whether causality runs from or toward the proximal construct, respectively.

if the measure refers to a cause or an effect of the construct, their relationship is indirect, spurious, or unanalyzed.

Naturally, the models in Table 1 do not exhaust all possible causal structures. For example, the indirect reflective and indirect formative models can be elaborated by including additional mediators of the relationship between the construct and its measures. The models can also be combined to form hybrid models. For instance, the direct reflective and direct formative models can be combined to depict reciprocal causation between a construct and its measures (Bollen, 1989) or to incorporate both reflective and formative measures of a single construct (Jöreskog & Goldberger, 1975). Thus, the models in Table 1 represent conceptual building blocks that can be used to explicate the relationship between constructs and measures.

Illustrative Applications

The preceding discussion developed principles for specifying the direction and structure of relationships between constructs and measures. For expository purposes, we applied selected aspects of these principles to various constructs and measures (e.g., SES, involvement, satisfaction, performance). We now more fully apply these principles to constructs and measures characteristic of psychological, sociological, and organizational research.

Life Stress

Following Selye (1956), stress researchers often conceptualize life stress as the degree of change required by major life events (Dohrenwend & Dohrenwend, 1981). The most widely used measure of life events is the 43-item Social Readjustment Rating Scale (SRRS), developed by Holmes and Rahe (1967) to measure "the amount and duration of change in one's accustomed pattern of life" (p. 213). Studies of the factor structure and the internal consistency reliability of the SRRS (e.g., Kipper & Furcon, 1981; Lei & Skinner, 1980) imply that the relationship between the SRRS and life stress follows the direct reflective model. However, Turner and Wheaton (1995) argued that life event measures such as the SRRS should not exhibit internal consistency because items in these measures are not alternative indicators of a single underlying construct but instead describe distinct events a person might experience. On the basis of similar reasoning, Cohen et al. (1990) suggested that life

event measures can be viewed as causes of life stress, corresponding to the direct formative model.

The preceding controversy can be resolved by applying principles developed in this article to the SRRS (see Table 2). Specifically, the life stress construct (i.e., change in life pattern) and the SRRS are distinct entities, given that life change is distinct from ratings of events that may prompt change. Association is plausible, although change in life pattern may occur for reasons other than the 43 events that constitute the SRRS. Temporal precedence is likewise difficult to establish, given that SRRS scores should logically follow the occurrence of their corresponding events but can follow or precede the change in life pattern these events presumably create. These ambiguities are symptomatic of a rival causal explanation for the relationship between life change and the SRRS in which the occurrence of individual life events causes change in life pattern and produces scores on the SRRS. The association between event occurrence and the SRRS scores is imperfect because of reporting biases and errors in recall (Raphael, Cloitre, & Dohrenwend, 1991). On the basis of this reasoning, the relationship between life change and the SRRS would follow the spurious model with multiple common causes (Figure 5B), in which the common causes are individual life events.

Organizational Commitment

Mowday, Steers, and Porter (1979) defined organizational commitment as "the relative strength of an individual's identification with and involvement in a particular organization" (p. 226). To measure this construct, Mowday et al. developed the 15-item Organizational Commitment Questionnaire (OCQ). Although Mowday et al. stated that commitment may consist of multiple factors, they evaluated the OCQ using internal consistency reliability estimation and common factor analysis and concluded that the OCQ represents "a single common underlying construct" (p. 232). Subsequent studies have treated the OCQ as a unidimensional measure and have routinely reported internal consistency reliability estimates (Mathieu & Zajac, 1990). Hence, the development and application of the OCQ have implicitly followed the direct reflective model.

Closer scrutiny of the OCQ suggests that the direct reflective model may be inappropriate (see Table 2). For illustration, we focus on OCQ Items 5 ("I find that my values and the organization's values are very similar"), 14 ("For me this is the best of all possible or-

Table 2
Application of Guidelines to Illustrative Measures

Construct	Instrument	Current model	Causal direction					
			Distinct entity	Association	Temporal precedence	Rival causal explanation	Causal structure	Revised model
Life stress	SRRS (Holmes & Rahe, 1967)	Either direct reflective or direct formative	Change in life pattern distinct from scores	Probable, but life change may occur for reasons not described by SRRS items	Life stress is experienced before scores are reported	Experience of individual events causes scores and life stress	Both scores and life stress caused by life event constructs	Spurious
Organizational commitment	OCQ (Mowday, Steers, & Porter, 1979)	Direct reflective	Commitment distinct from scores	Commitment and scores on most items should covary	Commitment is experienced before scores are reported	Scores reflect constructs that are causes and consequences of commitment	Commitment has indirect effects on attitudes and behaviors	Combined spurious and indirect reflective
Social interaction	7-item measure (Doney & Cannon, 1997)	Direct formative	Interaction distinct from scores	Interaction and scores should covary	Interaction occurs before scores are reported	Specific types of interaction cause scores	Types of interaction directly affect scores	Direct reflective with several specific interaction constructs

Note. SRRS = Social Readjustment Rating Scale; OCQ = Organizational Commitment Questionnaire.

ganizations for which to work"), and 2 ("I talk up this organization to my friends as a great place to work for"). Scores on these items are distinct from the commitment construct, which appears to represent an attitude within the mind of the respondent (assuming "involvement" refers to the importance ascribed to the organization; cf. Kanungo, 1982). The commitment construct is likely to precede the OCQ item scores, assuming respondents have some notion of their commitment before the OCQ is administered. Association is plausible in that higher commitment should logically be associated with higher scores on the three illustrative OCQ items. However, the relationship between the commitment construct and the OCQ scores may follow a causal structure more complex than the direct reflective model. In particular, Item 5 may be viewed as a direct reflective measure of a cause of commitment, in that congruence between personal and organizational values may create a sense of identification with the organization (Meglino, Ravlin, & Adkins, 1989). Thus, the relationship between Item 5 and commitment may follow the spurious model (Figure 5A). Conversely, Items 14 and 2 may be considered direct reflective measures of evaluative and behavioral consequences of commitment, respectively, in that identification with an organization should create positive appraisals of that organization and prompt the person to speak positively about the organization. Thus, the relationship of commitment with Items 14 and 2 follows the indirect reflective model. Combining these models indicates that the relationship between organizational commitment and OCQ Items 5, 14, and 2 is captured not by the direct reflective model but rather by a combined spurious and indirect reflective model. Similar reasoning can be applied to the remaining 12 OCQ items to develop a full model of their relationship with the commitment construct.

Social Interaction

In a study of buyer–seller relationships, Doney and Cannon (1997) conceptualized social interaction as contact between a buyer and a seller in an informal environment conducive to enhanced information flow. To measure social interaction, Doney and Cannon used seven items describing various forms of social interaction (e.g., meeting away from the workplace, talking about common interests). Because the forms of interaction differed, Doney and Cannon considered this measure formative, thereby adopting the direct formative model.

Further evaluation of Doney and Cannon's (1997)

measure suggests a different model. Scores on items describing social interaction are distinct from those interactions, and higher scores should be associated with the increased occurrence of interactions. However, social interaction scores should follow the interactions they describe, given that the scores reflect events that have already occurred. This temporal ordering is the opposite of that implied by the direct formative model and indicates a rival causal explanation in which constructs signifying forms of social interaction generate scores on their corresponding items. Inspection of Doney and Cannon's items suggests that these constructs may include informal conversations ("talk about family, sports or other personal interests," "talk about common interests beside work") and informal meetings (i.e., "meet away from the workplace," "meet over breakfast, lunch or dinner," "attend entertainment events [sports, theater, etc.]," "get together primarily to have fun," "get together with other family members"). Effects of these constructs on their measures are presumably direct because the items describe the inherent attributes of their constructs. This reasoning yields two direct reflective models, one for informal conversations and another for informal meetings. The informal-meetings construct can be subdivided into constructs that describe specific types of meetings, thereby reflecting the distinctions among the items (e.g., an entertainment event may not entail a meal or contact with family members). If this degree of specificity is not desired, items describing specific types of meetings can be replaced with general items akin to "meet away from the workplace," yielding additional direct reflective measures of informal meetings. Note that this respecified model does not entail a *social interaction* construct, because this term merely serves as a label for a category that encompasses different ways of interacting informally. These forms of interaction can be investigated jointly to examine the general phenomenon of social interaction.

Summary and Implications

This article has derived models that describe relationships between constructs and measures and has provided guidelines for choosing among the models on the basis of principles for determining the causal direction and structure of relationships between constructs and measures. The models derived in this article go beyond simple direct effects specified in most measurement models to include indirect, spurious, and unanalyzed associations between constructs and

measures. Guidelines for choosing among the models apply conditions for causality from the philosophy of science literature, thereby supplementing the commonsense assertions often used to specify causal direction between constructs and measures. Throughout the article, we have underscored the premise that the relationship between a construct and its measures cannot be specified without reference to the conceptual underpinnings of the construct and measures. Thus, rather than generating narrow prescriptions, this article has attempted to provide a way of thinking about linkages between constructs and measures that will shed light on the nature and direction of their relationship.

We hope the guidelines offered in this article will help researchers specify relationships between constructs and measures with clarity and precision, thereby enhancing the quality and rigor of auxiliary theories linking constructs to measures. As auxiliary theories are improved, researchers can better rely on these theories to drive analytical decisions, as is widely recommended for tests of substantive theories (MacCallum, Roznowski, & Necowitz, 1992; Runkel & McGrath, 1972). Moreover, stronger auxiliary theories will help reduce misspecifications of relationships between constructs and measures. Such misspecification can lead to gross inaccuracies in tests of relationships among substantive constructs (MacCallum & Browne, 1993), inviting erroneous conclusions and thwarting the accumulation of knowledge. Thus, by improving auxiliary theories linking constructs to measures, tests of substantive theories relating constructs to one another will likewise improve.

On the basis of our experience with the guidelines developed in this article, we believe their application will yield three general conclusions, each of which has important substantive ramifications. First, many instances of the direct formative model can be respecified as the spurious model, as illustrated by our discussion of the measurement of SES. Researchers have noted that measures in the direct formative model are assumed to contain no measurement error (Heise, 1972; MacCallum & Browne, 1993), an assumption that is untenable in most situations. This assumption can be relaxed by incorporating constructs that each cause their own reflective measures and collectively cause the construct of interest. To identify all parameters in this model, direct reflective measures can be added to the construct of interest. This respecified model would provide estimates of relationships between each construct and its measures as well as

structural relationships between constructs. As pointed out long ago by Blalock (1971), "many so-called 'background variables' such as race, sex, occupation of father, religion, community, region, and even 'age' are basically crude indicators of stimuli that are thought to influence behavior" (p. 337). Consequently, these background variables should be specified not as error-free formative measures but as imperfect reflective measures of constructs that cause other constructs.

Second, items that describe distinct facets of a general construct are often misspecified as direct reflective measures of that construct. This tendency is illustrated by the measurement of overall job satisfaction by using items that describe satisfaction with job facets (Warr et al., 1979; Weiss et al., 1967). In general, facet measures should be respecified as reflective measures of facet constructs, which in turn (a) influence a general construct, as in the relationship between facet satisfaction and overall job satisfaction, or (b) are influenced by a superordinate construct, as in second-order factor models. In the former case, the relationship between the facet measures and the general construct would follow the spurious model shown in Figure 5B. In the latter case, the relationship between the facet measures and the superordinate construct would follow the indirect reflective model shown in Figure 3B. In neither case would facet measures serve as direct reflective measures of the general construct.

Third, researchers should refrain from invoking the direct formative model as an explanation for low reliability estimates. Computerized searches for studies citing two recent articles on formative measurement (Bollen & Lennox, 1991; MacCallum & Browne, 1993) revealed several instances in which these articles were used to construct post hoc explanations for low reliability estimates (e.g., Adkins, 1995; Jarley, Fiorito, & Delaney, 1997; Judge & Bretz, 1994; Thomas & Ravlin, 1995). As we have argued, the specification of measures as formative or reflective should be based on a priori conceptual criteria, not on post hoc empirical evidence, particularly when that evidence consists of a low reliability estimate. Furthermore, if measures are specified as formative, their validity must still be established. It is bad practice to report a low reliability estimate, claim that one's measures are formative, and do nothing more.

Conclusion

This article has advanced guidelines for specifying the causal direction and the causal structure of the

relationship between constructs and measures. By applying these guidelines, researchers can develop auxiliary theories linking constructs to measures with the rigor and clarity used to specify relationships among theoretical constructs. Thus, we advocate a holistic construal (Bagozzi & Phillips, 1982) in which auxiliary and substantive theories are developed jointly to enhance our understanding of relationships among theoretical constructs and the mapping of these constructs onto the empirical world.

References

Adkins, C. L. (1995). Previous work experience and organizational socialization: A longitudinal examination. *Academy of Management Journal, 38,* 839–862.

Aldag, R. J., & Brief, A. P. (1978). Examination of alternative models of job satisfaction. *Human Relations, 31,* 91–98.

Alwin, D. F., & Hauser, R. M. (1975). The decomposition of effects in path analysis. *American Sociological Review, 40,* 37–47.

Alwin, D. F., & Tessler, R. C. (1974). Causal models, unobserved variables, and experimental data. *American Journal of Sociology, 80,* 58–86.

Asher, H. B. (1983). *Causal modeling.* Newbury Park, CA: Sage.

Averill, J. R. (1973). Personal control over aversive stimuli and its relationship to stress. *Psychological Bulletin, 80,* 286–303.

Bagozzi, R. P. (1977). Structural equation models in experimental research. *Journal of Marketing Research, 14,* 209–226.

Bagozzi, R. P. (1980). *Causal models in marketing.* New York: Wiley.

Bagozzi, R. P., & Fornell, C. (1982). Theoretical concepts, measurements, and meaning. In C. Fornell (Ed.), *A second generation of multivariate analysis: Vol. 2. Measurement and evaluation* (pp. 24–38). New York: Praeger.

Bagozzi, R. P., & Phillips, L. W. (1982). Representing and testing organizational theories: A holistic construal. *Administrative Science Quarterly, 27,* 459–489.

Bagozzi, R. P., & Yi, Y. (1990). Assessing method variance in multitrait–multimethod matrices: The case of self-reported affect and perceptions at work. *Journal of Applied Psychology, 75,* 547–560.

Bedeian, A. G., Day, D. V., & Kelloway, E. K. (1997). Correcting for measurement error attenuation in structural equation models: Some important reminders. *Educational and Psychological Measurement, 57,* 785–799.

Binning, J. F., Zaba, A. J., & Whattam, J. C. (1986). Ex-

plaining the biasing effects of performance cues in terms of cognitive categorization. *Academy of Management Journal, 29,* 521–535.

Blalock, H. M. (1964). *Causal inferences in nonexperimental research.* Chapel Hill: University of North Carolina Press.

Blalock, H. M. (1971). Causal models involving unobserved variables in stimulus–response situations. In H. M. Blalock (Ed.), *Causal models in the social sciences* (pp. 335–347). Chicago: Aldine.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

Bollen, K. A., & Davis, W. R. (1994, August). *Causal indicator models: Identification, estimation, and testing.* Paper presented at the 1993 American Sociological Association Convention, Miami, FL.

Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin, 110,* 305–314.

Bollen, K. A., & Ting, K. (1993). Confirmatory tetrad analysis. In P. V. Marsden (Ed.), *Sociological methodology 1993* (pp. 147–175). Washington, DC: American Sociological Association.

Brief, A. P., Butcher, A. H., & Roberson, L. (1995). Cookies, disposition, and job attitudes: The effects of positive mood-inducing events and negative affectivity on job satisfaction in a field experiment. *Organizational Behavior and Human Decision Processes, 62,* 55–62.

Burt, R. S. (1976). Interpretational confounding of unobserved variables in structural equation models. *Sociological Methods and Research, 5,* 3–52.

Campbell, D. T. (1960). Recommendations for APA test standards regarding construct, trait, or discriminant validity. *American Psychologist, 15,* 546–553.

Cohen, P., Cohen, J., Teresi, J., Marchi, M., & Velez, C. N. (1990). Problems in the measurement of latent variables in structural equations causal models. *Applied Psychological Measurement, 14,* 183–196.

Coleman, J. S. (1976). Regression analysis for the comparison of school and home effects. *Social Science Research, 5,* 1–20.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Boston: Houghton Mifflin.

Coovert, M. D., Craiger, J. P., & Teachout, M. S. (1997). Effectiveness of the direct product versus confirmatory factor model for reflecting the structure of multimethod–multirater job performance data. *Journal of Applied Psychology, 82,* 271–280.

Costner, H. L. (1969). Theory, deduction, and the rules of

correspondence. *American Journal of Sociology, 75,* 245–263.

Costner, H. L. (1971). Utilizing causal models to discover flaws in experiments. *Sociometry, 34,* 398–410.

Cronbach, L. J., & Meehl, P. C. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281–302.

Delanty, G. (1997). *Social science: Beyond constructivism and realism.* Minneapolis: University of Minnesota Press.

DeVellis, R. F. (1991). *Scale development: Theories and applications.* Newbury Park, CA: Sage.

Dohrenwend, B. S., & Dohrenwend, B. P. (1981). *Stressful life events and their contexts.* New York: Neale Watson.

Doney, P. M., & Cannon, J. P. (1997). An examination of the nature of trust in buyer–seller relationships. *Journal of Marketing, 61,* 35–51.

Duncan, O. D. (1975). *Introduction to structural equation models.* New York: Academic Press.

Feldman, J. M., & Lynch, J. G., Jr. (1988). Self-generated validity and other effects of measurement of belief, attitude, intention, and behavior. *Journal of Applied Psychology, 73,* 421–435.

Ferratt, T. W. (1981). Overall job satisfaction: Is it a linear function of facet satisfaction? *Human Relations, 34,* 463–473.

Fornell, C., & Bookstein, F. L. (1982). Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. *Journal of Marketing Research, 19,* 440–452.

Fox, J. (1980). Effect analysis in structural equation models. *Sociological Methods and Research, 9,* 3–28.

Garthwaite, P. H. (1994). An interpretation of partial least squares. *Journal of the American Statistical Association, 89,* 122–127.

Hackman, J. R., & Oldham, G. R. (1980). *Work redesign.* Reading, MA: Addison-Wesley.

Harman, H. H. (1976). *Modern factor analysis* (3rd ed.). Chicago: University of Chicago Press.

Hauser, R. M. (1972). Disaggregating a social-psychological model of educational attainment. *Social Science Research, 1,* 159–188.

Hauser, R. M., & Goldberger, A. S. (1971). The treatment of unobservable variables in path analysis. In H. L. Costner (Ed.), *Sociological methodology 1971* (pp. 81–117). San Francisco: Jossey-Bass.

Heise, D. R. (1972). Employing nominal variables, induced variables, and block variables in path analysis. *Sociological Methods & Research, 1,* 147–173.

Heise, D. R. (1975). *Causal analysis.* New York: Wiley.

Holmes, T. H., & Rahe, R. H. (1967). The Social Readjust-

ment Rating Scale. *Journal of Psychosomatic Research, 11,* 213–218.

Hume, D. (1946). *Inquiries concerning the human understanding.* London: Oxford University Press.

Igra, A. (1979). On forming variable set composites to summarize a block recursive model. *Social Science Research, 8,* 253–264.

Ironson, G. H., Smith, P. C., Brannick, M. T., Gibson, W. M., & Paul, K. B. (1989). Construction of a Job in General scale: A comparison of global, composite, and specific measures. *Journal of Applied Psychology, 74,* 193–200.

James, L. R., Mulaik, S. A., & Brett, J. M. (1982). *Causal analysis: Assumptions, models and data.* Beverly Hills, CA: Sage.

Jarley, P., Fiorito, J., & Delaney, J. T. (1997). A structural contingency approach to bureaucracy and democracy in U.S. national unions. *Academy of Management Journal, 40,* 831–861.

Johns, G. (1994). How often were you absent? A review of the use of self-reported absence data. *Journal of Applied Psychology, 79,* 574–591.

Jöreskog, K. G., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association, 10,* 631–639.

Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: User's reference guide.* Chicago: Scientific Software International.

Judge, T. A., & Bretz, R. D. (1994). Political influence behavior and career success. *Journal of Management, 20,* 43–65.

Kanungo, R. N. (1982). Measurement of job and work involvement. *Journal of Applied Psychology, 67,* 341–349.

Kim, J. O., & Mueller, C. W. (1978). *Factor analysis.* Beverly Hills, CA: Sage.

Kipper, D. A., & Furcon, J. (1981). Factors underlying the Social Readjustment Rating Scale. *Personality and Individual Differences, 2,* 31–36.

Lee, S., & Hershberger, S. (1990). A simple rule for generating equivalent models in covariance structure modeling. *Multivariate Behavioral Research, 25,* 313–334.

Lei, H., & Skinner, H. A. (1980). A psychometric study of life events and social readjustment. *Journal of Psychosomatic Research, 24,* 57–65.

Locke, E. A. (1976). The nature and causes of job satisfaction. In M. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 1297–1350). Chicago: Rand McNally.

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports, 3,* 635–694.

Long, J. S. (1983). *Confirmatory factor analysis: A preface to LISREL.* Beverly Hills, CA: Sage.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

MacCallum, R. C., & Browne, M. W. (1993). The use of causal indicators in covariance structure models: Some practical issues. *Psychological Bulletin, 114,* 533–541.

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modification in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin, 111,* 490–504.

MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin, 114,* 185–199.

Marsden, P. V. (1982). A note on block variables in multiequation models. *Social Science Research, 11,* 127–140.

Mathieu, J. E., & Zajac, D. M. (1990). A review and meta-analysis of the antecedents, correlates, and consequences of organizational commitment. *Psychological Bulletin, 108,* 171–194.

Meglino, B. M., Ravlin, E. C., & Adkins, C. L. (1989). A work values approach to corporate culture: A field test of the value congruence process and its relationship to individual outcomes. *Journal of Applied Psychology, 74,* 424–434.

Messick, S. (1981). Constructs and their vicissitudes in educational and psychological measurement. *Psychological Bulletin, 89,* 575–588.

Messick, S. (1995). Validity of psychological assessment. *American Psychologist, 50,* 741–749.

Mill, J. S. (1886). *A system of logic, ratiocinative and inductive* (8th ed.). London: Oxford University Press.

Mobley, W. H. (1982). *Employee turnover: Causes, consequences, and control.* Reading, MA: Addison-Wesley.

Mowday, R. T., Steers, R. M., & Porter, L. W. (1979). The measurement of organizational commitment. *Journal of Vocational Behavior, 14,* 224–247.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed). New York: McGraw-Hill.

Phillips, J. S., & Lord, R. G. (1982). Schematic information processing and perceptions of leadership in problem-solving groups. *Journal of Applied Psychology, 67,* 486–492.

Popper, K. R. (1959). *The logic of scientific discovery.* New York: Basic Books.

Raphael, K. G., Cloitre, M., & Dohrenwend, B. P. (1991). Problems of recall and misclassification with checklist methods of measuring stressful life events. *Health Psychology, 10,* 62–74.

Rindskopf, D., & Rose, T. (1988). Some theory and application of confirmatory second-order factor analysis. *Multivariate Behavioral Research, 23,* 51–67.

Runkel, P. J., & McGrath, J. E. (1972). *Research on human behavior: A systematic guide to method.* New York: Holt.

Schwab, D. P. (1980). Construct validity in organizational behavior. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior* (Vol. 2, pp. 3–43). Greenwich, CT: JAI Press.

Selye, H. (1956). *The stress of life.* New York: McGraw-Hill.

Suppes, P. (1970). *A probabilistic theory of causation.* Amsterdam: North-Holland.

Thomas, D. C., & Ravlin, E. C. (1995). Responses of employees to cultural adaptation by a foreign manager. *Journal of Applied Psychology, 80,* 133–146.

Thompson, B. (1984). *Canonical correlation analysis: Uses and interpretation.* Newbury Park, CA: Sage.

Turner, R. J., & Wheaton, B. (1995). Checklist measurement of stressful life events. In S. Cohen, R. C. Kessler, & L. U. Gordon (Eds.), *Measuring stress: A guide for health and social scientists* (pp. 29–58). Oxford, England: Oxford University Press.

Warr, P. B., Cook, J. D., & Wall, T. D. (1979). Scales for the measurement of some work attitudes and aspects of psychological well-being. *Journal of Occupational Psychology, 52,* 129–148.

Weick, K. A. (1989). Theory construction as disciplined imagination. *Academy of Management Review, 14,* 516–531.

Weiss, D. J., Dawis, R. W., England, G. W., & Lofquist, L. H. (1967). *Manual for the Minnesota Satisfaction Questionnaire.* Minneapolis: University of Minnesota, Industrial Relations Center.

Williams, L. J., & Brown, B. K. (1994). Method variance in organizational behavior and human resources research: Effects on correlations, path coefficients, and hypothesis testing. *Organizational Behavior and Human Decision Processes, 57,* 185–209.

Wold, H. (1982). Soft modeling: The basic design and some extensions. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observation* (Part 2, pp. 1–54). Amsterdam: North-Holland.

Zuriff, G. (1998). Against metaphysical social constructionism in psychology. *Behavior and Philosophy, 26,* 5–28.