Psychological Methods
2000, Vol. 5, No. 1, 3–22

1082-989X/00/$5.00 DOI: 10.1037//1082-989X.5.1.3

A Tetrad Test for Causal Indicators

Kenneth A. Bollen
University of North Carolina at Chapel Hill

Kwok-fai Ting
The Chinese University of Hong Kong

The authors propose a confirmatory tetrad analysis test to distinguish causal from effect indicators in structural equation models. The test uses "nested" vanishing tetrads that are often implied when comparing causal and effect indicator models. The authors present typical models that researchers can use to determine the vanishing tetrads for 4 or more variables. They also provide the vanishing tetrads for mixtures of causal and effect indicators, for models with fewer than 4 indicators per latent variable, or for cases with correlated errors. The authors illustrate the test results for several simulation and empirical examples and emphasize that their technique is a theory-testing rather than a model-generating approach. They also review limitations of the procedure including the indistinguishable tetrad equivalent models, the largely unknown finite sample behavior of the test statistic, and the inability of any procedure to fully validate a model specification.

Nearly all treatments of measurement in the social sciences treat observed variables as dependent on latent variables. According to this view, a shift in the construct leads to an expected shift in an indicator. Following Blalock (1964), we refer to these measures as *effect indicators*. For instance, the responses to a series of items on a math test should reflect a student's quantitative ability. A person's degree of agreement with questions about whether an individual is "as good as others" or whether the individual has "pride in self" are likely to be effect indicators of self-esteem. This effect indicator perspective underlies factor analysis (Spearman, 1904) and classical test theory (Lord & Novick, 1968). It is the basis for most measures of reliability and validity that are common in psychology and the social sciences.

Effect indicators are appropriate for many situations in psychological measurement, but they are not appropriate for all situations. Several researchers (Blalock, 1964; Bollen, 1984, 1989; Bollen & Lennox, 1991; Hayduk, 1987; Land, 1970; MacCallum & Browne, 1993) have noted that some observed variables are more appropriately treated as determinants rather than effects of the latent variable. Blalock (1964) called these *causal indicators*. For instance, loss of job, divorce, or birth of a child are measures of exposure to stress that are best thought of as causal indicators. That is, each event is a determinant of exposure to stress rather than a consequence of it. A social psychologist interested in the degree of social interaction should treat time spent with family, time spent with friends, and time spent with coworkers as causal indicators of the amount of social interaction. Quality of life might be gauged by indicators such as self-reported health, happiness, and economic status, but it is doubtful that we can treat these as effect indicators. A psychologist studying accuracy of memory might indicate the number of details correctly recalled, but each detail could be thought of as a causal indicator of the latent variable of the construct of memory accuracy. In field research, it might be necessary to control for socioeconomic status of an individual. Education, income, and occupational prestige are likely to be causal indicators of SES. In other cases the indicators of a psychological construct might be a mixture of effect and causal indicators. For instance, Bollen and Lennox (1991, p. 311) suggested that the Center for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977) has some effect

Kenneth A. Bollen, Department of Sociology, University of North Carolina at Chapel Hill; Kwok-fai Ting, Department of Sociology, The Chinese University of Hong Kong, Shatin, Hong Kong, China.

Partial support for this work came from the Sociology Program of the National Science Foundation (SES-9121564). We appreciate the comments of Walter Davis.

Correspondence concerning this article should be addressed to Kenneth A. Bollen, CB 3210 Hamilton, Department of Sociology, University of North Carolina, Chapel Hill, North Carolina 27599-3210. Electronic mail may be sent to bollen@unc.edu.

indicators (e.g., "I felt depressed" and "I felt sad") mixed with causal indicators (e.g., "I felt lonely").

Though we make no claim that most measures in psychology are really causal rather than effect indicators, we do claim that causal indicators are sometimes present but rarely considered.[1] Knowledge of whether observed variables are causal or effect indicators is important for several reasons. First, causal (or *formative*) indicators have different measurement properties than effect indicators. For instance, whereas effect indicators positively associated with the same unidimensional latent variable should be positively correlated, the same is not true for causal indicators (Bollen, 1984; Bollen & Lennox, 1991). Factor analysis and many of the techniques of index construction and measurement validation are based on the implicit assumption of effect indicators. Because of this, observed variables that are causal indicators of a latent variable may be incorrectly discarded as invalid measures. Learning that a variable is a determinant rather than an effect of the latent variable could prevent this incorrect decision. Furthermore, with effect indicators we can sample measures of a latent variable. With causal indicators a more exhaustive list of indicators should be included. The sum of the omitted causal indicators should be uncorrelated with the included ones (see Bollen & Lennox, 1991, p. 308).

Second, wrongly treating variables as effect rather than causal indicators leads to model specification error. The specification error can bias parameter estimators and lead to incorrect assessments of the relationships between variables. Even if a researcher's main interest is in the latent variable model, the bias from the treatment of causal indicators as effect indicators can spread to parameters outside of the measurement model. A related problem is that such a misspecification distorts the researcher's understanding of the operation of the system. For instance, an effect indicator is changed by shifts in the latent variable. The effect indicator's main value is in providing a way to track the progress of the construct. A causal indicator is an observed variable which, rather than tracks, helps to determine the path of the latent variable. In other words, manipulation of the causal indicator is a way to indirectly manipulate the latent variable, though other variables and a disturbance also affect the latent variable. Ideally, to fully validate the usefulness of the latent variable or construct that is influenced by causal indicators, there should be other latent or observed variables that the construct influences so that we can assess whether the construct behaves as hypothesized. One model that does this is the multiple indicator and multiple causes (MIMIC) model (Hauser & Goldberger, 1971) that has both causal and effect indicators of the latent variable.

Establishing the causal priority between a latent variable and its indicators can be difficult. One aid is to perform "mental experiments," in which a researcher imagines a shift in the latent variable and then judges whether a simultaneous shift in all the observed variables is likely. If so, then this is consistent with an effect indicator specification. Alternatively, if the researcher imagines a shift in an observed variable as leading to a shift in the latent variable even if there is no change in the other indicators, then this is consistent with a causal indicator model (Bollen, 1989, pp. 65–67). These mental experiments take advantage of a researcher's understanding of a substantive area to help order the relation between the latent and observed variables. However, the results can be ambiguous with no clear resolution. It also does not provide an empirical means to check the specification.

In some rare cases it is possible to devise experiments that help to test whether variables are causal or effect indicators (Bollen, 1989, pp. 66–67), however this will be very difficult in most practical situations. Estimating two models, one with causal indicators and another with effect indicators, does not solve the problem. The parameters of one model are not a more restrictive form of the parameters in another model, so we cannot turn to the traditional likelihood ratio test to compare their fit. Also, a model that has a latent variable with only causal indicators is sometimes underidentified, which creates difficulties in estimation (Bollen, 1989, pp. 312–313; MacCallum & Browne, 1993).

Our article has several purposes. First, we develop a vanishing tetrad test that provides an empirical test of whether a causal or effect indicator specification is appropriate. Second, we illustrate the test with several empirical examples. Finally, we discuss the merits and limitations of our test. To avoid any misunderstandings, we emphasize that our task is to provide a test of two or more well-formulated models, one with causal indicators and the other with effect indicators.

[1] See Ozer and Reise (1994, pp. 363–64) and Neuberg, West, Judice, and Thompson (1997) for examples of articles that discuss causal indicators in research on personality assessment.

We are not proposing a model-generating procedure that will create models that are consistent with data. Exploratory techniques that make use of vanishing tetrads, partial correlations, and graphical theory methods to develop models are available in the work of Glymour et al. (1987), Spirtes, Glymour, and Scheines (1993), and others. We will contrast these exploratory tetrad and other analysis techniques with the confirmatory tetrad analysis (CTA) that we use in the conclusion.

The next section of the article establishes the basic principles of determining vanishing tetrads for models whose causal and effect indicators have a factor complexity of one, that is, an observed variable serves as an indicator for no more than one factor and presents the test statistic. Our discussion covers complications to the basic model including the correlation between measurement errors, a mixture of causal and effect indicators, models with fewer than four indicators, causal indicators that contain random measurement error, and zero factor loadings. We present a simulation example and four empirical examples in the subsequent section. The concluding section highlights the limitations as well as the potential contributions of the test.

Vanishing Tetrads in Measurement Models

Tetrad refers to the difference between the product of a pair of covariances and the product of another pair among four random variables. For a foursome of variables, we can arrange the six covariances into three tetrads:

$$\tau_{1234} = \sigma_{12}\sigma_{34} - \sigma_{13}\sigma_{24},$$
$$\tau_{1342} = \sigma_{13}\sigma_{42} - \sigma_{14}\sigma_{32},$$

and

$$\tau_{1423} = \sigma_{14}\sigma_{23} - \sigma_{12}\sigma_{43}. \quad (1)$$

We use Kelley's (1928) notation that τ_{ghij} refers to $\sigma_{gh}\sigma_{ij} - \sigma_{gi}\sigma_{hj}$ and that σ is the population covariance of the two variables that are indexed below it. A vanishing tetrad means $\tau_{ghij} = 0$. The equations in Equation 1 apply to correlation data too, as they are only special cases of covariances where variables are rescaled to have a variance of one. Bollen and Ting (1993) provided test statistics that accommodate the test of vanishing tetrads for both correlation and covariance data.

Spearman (1904) introduced the idea of model-implied vanishing tetrads. Many researchers since him have explored the vanishing tetrads implied by a variety of factor analysis and more general structural equation models (SEMs; e.g., Glymour et al., 1987; Scheines, Spirtes, Glymour, & Meek, 1994; Spirtes, Glymour, & Scheines, 1993). Bollen and Ting (1993) proposed a CTA that tests one or several specific models. CTA is "confirmatory" in that models are specified in advance. The structure of each model often implies population tetrads that should be zero. A test of a model's vanishing tetrads is a test of the model's fit. Significant nonzero tetrads for the model-implied vanishing tetrads cast doubt on the appropriateness of the model. CTA also allows researchers to compare vanishing-tetrad nested models. Bollen and Ting (1998) further examined the asymptotic properties of the CTA test and proposed a bootstrap procedure in situations where the distribution of the test statistic may depart from the usual asymptotic one. With the computation program provided by Ting (1995), the CTA test is now an accessible procedure for researchers to test hypothesized models in terms of vanishing tetrads.[2] We follow Bollen and Ting's (1993) confirmatory strategy and apply CTA to help distinguish between causal and effect indicators. The procedure is most helpful when the researcher has narrowed down the plausible structures to a limited number of alternatives. We propose these tests for model testing, not as an exploratory tool to develop models.

There are several steps in this tetrad approach: (a) Specify the most plausible models of the relations between indicators and latent variables, (b) identify the model-implied vanishing tetrads for each model, (c) eliminate redundant vanishing tetrads, and (d) perform a simultaneous vanishing tetrad test. If the test indicates that effect indicators are plausible, we recommend that the researcher use conventional structural equation methods to estimate the structural parameters as a means to establish the statistical significance of the coefficients and variances. This last step is necessary because of the special cases where causal indicators have near zero covariances, and this leads to the vanishing tetrads. These four steps can be used on each latent variable in a model,

[2] The CTA-SAS program, which can be downloaded from http://www.cuhk.edu.hk/soc/ting/, is a confirmatory model-testing tool. It performs statistical tests on one or more sample vanishing tetrads simultaneously. In addition, it automatically handles redundancy problems among a set of vanishing tetrads and performs tests between nested models. See Ting (1995) for the full description of the program.

one at a time, in order to ascertain the accuracy of the measurement model. The test on each measurement submodel will only include the covariances among indicators of that submodel. We go over these steps in more detail in the sections that follow.

Causal and Effect Indicator Models

We assume that a researcher interested in applying our procedure has narrowed down the plausible model structures on the basis of substantive knowledge and prior research. The plausible structures are ones that differ in whether some of the indicators are determinants or consequences of latent variables. To compare causal indicator and effect indicator models, we need to determine the vanishing tetrads implied by these two types of models respectively. This can be done by means of covariance algebra. We illustrate this method for the basic models with four indicators and extend our discussions to models with more indicators.

Effect indicator model. Figure 1a is a model of a latent variable with four effect indicators. The equations corresponding to this diagram are

$$x_1 = \lambda_1\xi + \delta_1,$$
$$x_2 = \lambda_2\xi + \delta_2,$$
$$x_3 = \lambda_3\xi + \delta_3,$$

and

$$x_4 = \lambda_4\xi + \delta_4, \quad (2)$$

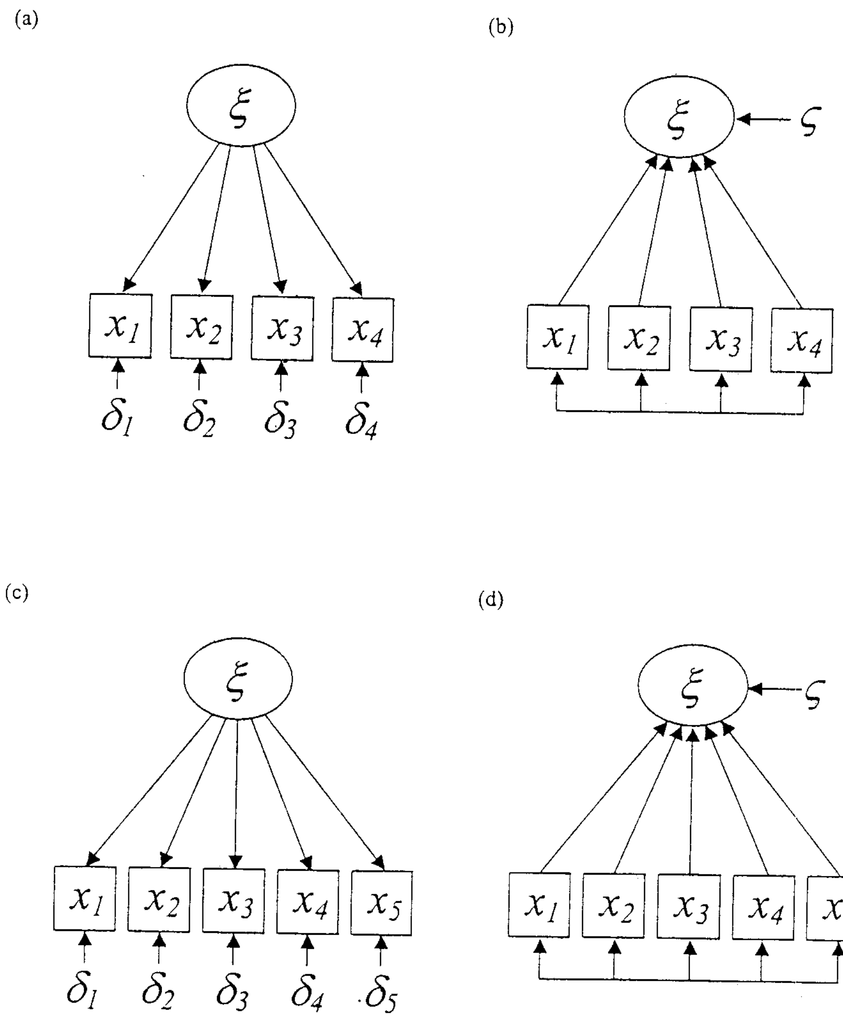

Figure 1. Four measurement models. The effect indicator models in (a) and (c) imply that all tetrads are vanished, whereas the causal indicator models in (b) and (d) entail no vanishing tetrads.

where δ_i is the random measurement error term with $E(\delta_i) = 0$ for all i, $\mathrm{COV}(\delta_i, \delta_j) = 0$ for $i \neq j$, and $\mathrm{COV}(\xi,\delta_i) = 0$ for all i. This means that we assume that all error variables have means of zero, are uncorrelated with each other, and are uncorrelated with the latent variable. All variables are written as deviations from their means to simplify algebra. The population covariance between x_i and x_j is a function of the path coefficients, λ_i, and the variance of the latent variable, ϕ, in an effect indicator model:

$$\sigma_{ij} = \lambda_i\lambda_j\phi. \qquad (3)$$

In this example, we can construct three tetrad equations from the covariances among the four observed variables. A simple algebraic manipulation shows that the following equalities must hold:

$$\tau_{1234} = \sigma_{12}\sigma_{34} - \sigma_{13}\sigma_{24} = 0,$$
$$\tau_{1342} = \sigma_{13}\sigma_{42} - \sigma_{14}\sigma_{32} = 0,$$

and

$$\tau_{1423} = \sigma_{14}\sigma_{23} - \sigma_{12}\sigma_{43} = 0, \qquad (4)$$

where τ_{ghij} is the population tetrad difference. In this case, both $\sigma_{gh}\sigma_{ij}$ and $\sigma_{gi}\sigma_{hj}$ equal $\lambda_1\lambda_2\lambda_3\lambda_4\phi^2$, and the difference between them is zero for all three tetrad equations. Thus, three vanishing tetrads are implied by the effect indicator model in Figure 1a. Notice that the above equalities hold regardless of the values of the path coefficients and the variance of the latent variable. This shows that vanishing tetrads are determined by the structure, not the parameters, of a model.

Causal indicator model. In Figure 1b, the arrows are reversed, with the observed variables x_1 to x_4 influencing the latent variable so that $\xi = \gamma_1x_1 + \gamma_2x_2 + \gamma_3x_3 + \gamma_4x_4 + \zeta$. For this model, the disturbance, ζ, consists of all of the other variables that influence the latent variable, ξ, but that are not in the model. It is this disturbance that makes the latent variable, ξ, distinct from the simple linear combination of the causal indicators. The disturbance has a mean of zero, $E(\zeta) = 0$, and it is uncorrelated with the xs, $\mathrm{COV}(x_i, \zeta) = 0$. Typically, the xs will be associated as indicated by the two-headed arrows linking these variables in Figure 1b. However, these are unanalyzed associations for the causal indicator model.

Unlike the effect indicator model in Figure 1a, this causal indicator model is underidentified. Indeed, the meaning of the latent variable would be clearer if the model included two or more outcome variables that were influenced by the latent variable. Yet in the early stages of research, these outcome variables may not be available, and the researcher is more concerned with trying to distinguish whether the indicators are causal or effect indicators. As we show below, we can provide a test to distinguish between Figure 1a and 1b, despite the underidentification of the causal indicator model in Figure 1b.

The population covariance between x_i and x_j is

$$\sigma_{ij} = E(x_ix_j). \qquad (5)$$

Because the observed variables are exogenous, there are no constraints on the covariances among the causal indicators. Except in the unlikely circumstances that the values of $\sigma_{gh}\sigma_{ij}$ and $\sigma_{gi}\sigma_{hj}$ exactly cancel each other out, none of the tetrads in Equation 4 vanishes. There is, however, one particular instance in which $\sigma_{gh}\sigma_{ij}$ will equal $\sigma_{gi}\sigma_{hj}$ in a causal indicator model. This occurs when some of the causal indicators are not linearly related, that is, their covariances tend toward zero. If both sides of the tetrad difference have one or more covariances equal to zero, the tetrad vanishes. One check on this condition is a significance test of the null hypothesis that each covariance that appears in a tetrad is zero. Another check is to estimate an effect indicator model. If one or more of the path coefficients, λ_i, or the variance of the latent variable, ξ, is not significantly different from zero, then the causal indicator model is more plausible than the effect indicator one. This situation might occur if the causal indicators are a series of "random" events, each of which influences the latent variable. Holmes and Rahe's (1967) original discussion of life events would be an example of such random events. Such cases are easy to detect. We illustrate this case with a measurement model of stress in the example section.

Models with greater than four indicators. Although we began the discussion with four indicator models, their implications generalize to larger measurement models. With more indicators, we consider four variables at a time. For example, a five-indicator model has five different combinations of four variables, and each set of combinations has three tetrads. In general, there will be $n!/(n-4)!4!$ sets of tetrads for models with n observed indicators. Figure 1c shows a model of a latent variable with five effect indicators. We take out one variable each time to form a set of four variables. If we consider x_2 to x_5, for example, the causal structure in this set of variables is identical to that of Figure 1a, and the three implied vanishing tetrads are τ_{2345}, τ_{2453}, and τ_{2534}. The same causal structure is shared by the other sets of tetrads, and all of the 15 possible vanishing tetrads are implied in this model. Figure 1d can be considered in a similar fashion. Every set of four variables from the model shows an identical causal structure to that of Figure 1b, which implies no vanishing tetrads. In contrast to the effect indicator model in 1c, none of the 15 tetrads vanishes. The same results generalize to larger models with more indicators: Effect indicator models imply all possible vanishing tetrads, whereas causal indicator models imply none.

Other issues. In structural equation modeling, causal indicator models are often neglected partly because they are underidentified models without prior constraints on some of the unknown parameters. Furthermore, effect indicator models and causal indicator models are not nested in parameters and cannot be tested using the conventional likelihood ratio test. That is, the usual likelihood ratio test requires that the parameters of one model are a constrained form of those of another. The causal indicator model involves different parameters than does the effect indicator model and hence undermines the conventional likelihood ratio approach to nesting. In terms of vanishing tetrads, however, they are nested models because the vanishing tetrads implied by one model are a subset of those implied by another model. The effect indicator model implies all possible vanishing tetrads, whereas the causal indicator model implies none. A simultaneous test on whether the tetrads actually vanish will help to determine the plausibility of these competing structures.

One might think that whether we treat indicators as causal or effects of a latent variable is an arbitrary decision such that an analyst can always derive one model from the other. As such, testing for one model versus the other does not matter. We can illustrate the problem with this perspective with the simple four-indicator, one-latent-variable models in Figure 1, a and b. Suppose that the effect indicator model in Figure 1a (see Equation 2) is valid. We could manipulate Equation 2 such that we get a single equation with ξ on the left-hand side and the xs on the right-hand side, as in Equation 6 (see below).

$$\xi = (1/4)\,(\lambda_1^{-1}x_1 + \lambda_2^{-1}x_2 + \lambda_3^{-1}x_3 + \lambda_4^{-1}x_4 - \lambda_1^{-1}\delta_1 - \lambda_2^{-1}\delta_2 - \lambda_3^{-1}\delta_3 - \lambda_4^{-1}\delta_4). \qquad (6)$$

Although this looks like a causal indicator model in that the xs are "predictors" of ξ, there is an important distinction. Instead of having a disturbance, ζ, that is uncorrelated with the xs, we have a composite disturbance that is a function of the δs that are correlated with the xs. The latter correlation follows from the δs being the errors of measurements of the xs. Thus, the effect indicator model and causal indicator model are not interchangeable, and if we attempt to rewrite the effect indicator model as a causal indicator one, we are led to a model that violates the causal indicator model assumptions. Further, one should not be misled to believe that because we use a weighted sum of the observed variables to estimate factor scores in an effect indicator model, this weighted score is the same as a causal indicator model. The estimate of the factor score is not the same as the latent variable, and the factor score estimates derive from an effect indicator model. Though we do not show it here, we also cannot arbitrarily create a valid effect indicator model if the assumptions of the causal indicator model hold.

Complications in Measurement Models

In the previous section, we discussed measurement models with solely causal indicators or entirely effect indicators. Their implications for vanishing tetrads were easy to determine. There are variations, however, in these measurement models that may complicate the identification of vanishing tetrads. We consider four such cases: mixed indicator models, latent variables with fewer than four indicators, causal indicators that contain random measurement error, and correlations between measurement errors.

Mixed indicator models. In a measurement model, it is possible that some of the observed variables are effect indicators, whereas others should be treated as causal indicators. This type of model is referred to as the mixed indicator model. Perhaps the best known mixed model is the MIMIC model (Hauser & Goldberger, 1971; Jöreskog & Goldberger, 1975). Figure 2 lists measurement models with a mixture of four causal and effect indicators. Figure 2a has one causal indicator, and all three vanishing tetrads are implied in this model. The vanishing tetrads implied by Figure 2a are identical to those of Figure 1a, meaning that they are tetrad equivalent models and cannot be distinguished in terms of vanishing tetrads. It is interesting to note that these two models are also "equivalent models" in terms of the likelihood ratio test and will have identical fits to the data. An important lesson that generalizes from this example is that there are some tetrad equivalent models that differ in the composition of causal and effect indicators that we cannot distinguish with this tetrad test or with the usual like-

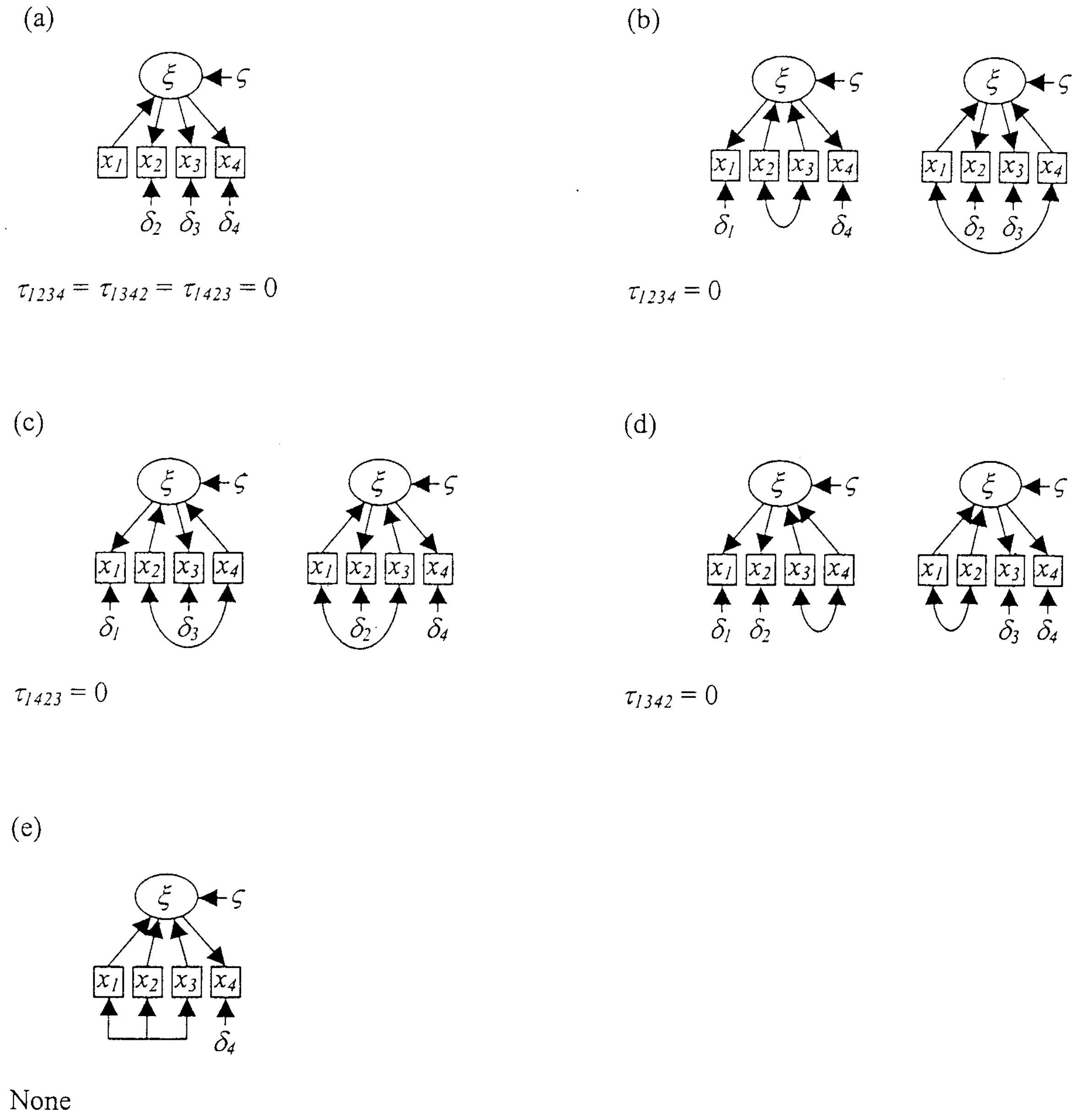

Figure 2. Measurement models with a mixture of causal and effect indicators. Below each path diagram are the implied vanishing tetrads for the model.

lihood ratio test. This is neither surprising nor new for those familiar with equivalent models in SEM (Frydenberg, 1990; Lee & Hershberger, 1990; Luijben, 1991; MacCallum, Wegener, Uchino, & Fabrigar, 1993; Verma & Pearl, 1990).

With two causal indicators, as shown in Figure 2, b–d, only one of the vanishing tetrads is implied. Among these six models, three pairs are tetrad equivalent models and each pair implies a different vanishing tetrad. Because these models are nested with models in Figure 1a and 2a, one can perform tests when competing conceptualizations involve these alternative models. Finally, there is no vanishing tetrad implied in Figure 2e that has three causal indicators. With no vanishing tetrad implied, any models with three causal indicators can be compared with those in Figure 1a and Figure 2, a–d.

Figure 2 also is useful in illustrating that not all mixed models will have nested vanishing tetrads. In the case where all indicators are either causal indicators or effect indicators, we can always consider them nested vanishing tetrads. In contrast, the mixed situation does not always result in nesting. For instance, if we wish to contrast the models in Figure 2, b and c, we cannot perform a nested tetrad test because each implies a different vanishing tetrad. However, we can compare either with the model in Figure 2a. Thus with mixtures of causal and effect indicators, we need to determine whether the vanishing tetrads are nested.

Using covariance algebra to identify the vanishing tetrads for mixed indicator models can be tedious when the number of indicators increases. Figure 2 can be helpful in these situations. No matter how big the measurement model is, any given four observed indicators fall into one of the model types in Figure 2. We illustrate this with an example of five indicators. Figure 3 shows a latent variable with x_1 and x_3 as causal indicators and x_2, x_4, x_5 as effect indicators. There are five sets of combinations, each composed of four different variables. These five sets of variables fall into two basic model types listed in Figure 2. Two sets of variables, x_1, x_2, x_4, x_5 and x_2, x_3, x_4, x_5, have an identical model structure as Figure 2a, which has one causal indicator. As a result, six vanishing tetrads, τ_{1245}, τ_{1452}, τ_{1542}, τ_{2345}, τ_{2453}, and τ_{2534}, are implied. The other three sets of variables—x_1, x_2, x_3, x_4, and x_1, x_3, x_4, x_5, and x_1, x_2, x_3, x_5—share one of the model structures from Figure 2, b–d, which has two causal and two effect indicators. With one vanishing tetrad implied in this model structure, three more vanishing tetrads, τ_{1423}, τ_{1453}, and τ_{1523}, are added to the model

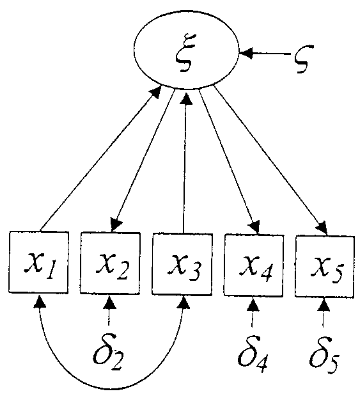

Figure 3. A five-indicator measurement model. This model has two causal indicators (x_1 and x_3) and three effect indicators (x_2, x_4, and x_5). Using Figure 2 and selecting four indicators at a time, one can determine the implied vanishing tetrads.

in Figure 3. For measurement models with a larger number of mixed causal and effect indicators, they always come down to the basic model types in Figure 2 when there are no correlated errors of measurement and the factor complexity of variables is no greater than one.

Latent variables with fewer than four indicators. The construction of tetrads requires at least four observed variables, and our discussion so far has been confined to latent variables with four or more indicators. Quite often we would like to test the causal relationship for a measurement model with fewer than four indicators. One feasible strategy is to take indicators from another latent variable to form a set of four variables. By doing so, we take into account the causal structure at the latent variable level to derive the implied constraints of the tetrad equations among the observed variables. Suppose we have only three indicators of a latent variable. We could use one indicator from another latent variable to evaluate whether we have causal or effect indicators. In the Appendix, we list the vanishing tetrads for a variety of models that assume one, two, or three of a latent variable's indicators are causal indicators. We do the same for a model where we have only two indicators for a latent variable, and we borrow two indicators from another latent variable to determine vanishing tetrads. Using Figure A1a of the Appendix as an example, we can determine the covariances among x_1 to x_4 by covariance algebra, which gives

$$\begin{aligned}\sigma_{12} &= \lambda_1\lambda_2\sigma_{\xi_1\xi_2},\\ \sigma_{13} &= \lambda_1\lambda_3\sigma_{\xi_1\xi_2},\\ \sigma_{14} &= \lambda_1\lambda_4\sigma_{\xi_1\xi_2},\\ \sigma_{23} &= \lambda_2\lambda_3\sigma_{\xi_2\xi_2},\\ \sigma_{24} &= \lambda_2\lambda_4\sigma_{\xi_2\xi_2},\end{aligned}$$

and

$$\sigma_{34} = \lambda_3\lambda_4\sigma_{\xi_2\xi_2}, \quad (7)$$

where λ_i is the path coefficient from the latent variable to x_i. It becomes clear that the product of two covariances, that is, $\sigma_{gh}\sigma_{ij}$, $\sigma_{gi}\sigma_{hj}$, and $\sigma_{gj}\sigma_{hi}$, in a tetrad is always equal to $\lambda_1\lambda_2\lambda_3\lambda_4\sigma_{\xi_1\xi_2}\sigma_{\xi_2\xi_2}$; thus, all three tetrads for x_1 to x_4 vanish. On the basis of the text and the figures in the Appendix, researchers can obtain vanishing tetrads for a wide variety of situations.

Correlated errors of measurement. All of the above discussion on measurement models is based on the assumption that measurement errors are uncorrelated. This assumption may be false. If the researcher specifies a model with correlations between measurement errors, the correlated errors will have consequences for the vanishing tetrads implied in a model. A vanishing tetrad, such as τ_{1234}, is implied only when $\sigma_{12}\sigma_{34} = \sigma_{13}\sigma_{24}$. If the measurement errors between x_3 and x_4, for example, are correlated, it adds a unique component, $COV(\delta_3, \delta_4)$, to the covariance term σ_{34}, making it very unlikely for the above equality to hold. This means a model with correlated errors implies fewer vanishing tetrads than the same model without correlated errors. In fact, vanishing tetrads of the former model are a subset of the latter model. To determine the vanishing tetrad implied by a measurement model with correlated error terms, one can first identify vanishing tetrads of the same model with no correlated error terms and then select the subset simply by removing those tetrads with covariances that involve correlated error terms.[3]

We demonstrate the above steps with a model in the Examples section. In the life satisfaction example, we test a model with five effect indicators from x_1 to x_5. This model has 15 tetrads, and all of them vanish. To further assess the power of the test, we examine the same model with the addition of a correlation between errors of x_3 and x_4 as an alternative model. The vanishing tetrads of this alternative model are a subset of the previous model without correlated errors. For those tetrads without the σ_{34} term, $\tau_{ghij} = 0$, because both $\sigma_{gh}\sigma_{ij}$ and $\sigma_{gi}\sigma_{hj}$ equal $\lambda_g\lambda_h\lambda_i\lambda_j\phi^2$. However, the following six tetrads that have σ_{34} in the equation will not vanish:

$$\begin{aligned}
\tau_{1234} &= \sigma_{12}\sigma_{34} - \sigma_{13}\sigma_{24} = \lambda_1\lambda_2\phi COV(\delta_3, \delta_4), \\
\tau_{1243} &= \sigma_{12}\sigma_{34} - \sigma_{14}\sigma_{23} = \lambda_1\lambda_2\phi COV(\delta_3, \delta_4), \\
\tau_{1354} &= \sigma_{13}\sigma_{45} - \sigma_{15}\sigma_{34} = \lambda_1\lambda_5\phi COV(\delta_3, \delta_4), \\
\tau_{1453} &= \sigma_{14}\sigma_{35} - \sigma_{15}\sigma_{34} = \lambda_1\lambda_5\phi COV(\delta_3, \delta_4), \\
\tau_{2354} &= \sigma_{23}\sigma_{45} - \sigma_{25}\sigma_{34} = \lambda_2\lambda_5\phi COV(\delta_3, \delta_4),
\end{aligned}$$

and

$$\tau_{2453} = \sigma_{24}\sigma_{35} - \sigma_{25}\sigma_{34} = \lambda_2\lambda_5\phi COV(\delta_3, \delta_4). \quad (8)$$

This is because $\sigma_{ij}\sigma_{34} = \lambda_i\lambda_j\lambda_3\lambda_4\phi^2 + \lambda_i\lambda_j\phi COV(\delta_3, \delta_4)$, and the second term on the right-hand side of the equation will not be cancelled out by either $\sigma_{i3}\sigma_{j4}$ or $\sigma_{i4}\sigma_{j3}$. Therefore, these tetrads in equation 8 should be excluded in the vanishing tetrad test. As a result, nine vanishing tetrads remain in the alternative model.

Causal indicators measured with error. In some situations the causal indicators are flawed measures of distinct substantive variables instead of indicators in the usual sense of the term. For instance, in an earlier example we referred to measures of education, income, and occupational prestige as causal indicators of the latent variable of socioeconomic status. It is likely that each indicator is an imperfect representation of the distinct substantive latent variables of education, income, and occupational prestige. If so, the paths to latent socioeconomic status should originate from the latent variables of education, income, and occupational prestige, rather than having direct paths from the corresponding indicators to socioeconomic status. Fortunately, the same vanishing tetrads occur whether we consider the causal indicators measured with random error, that is, $x_i = \lambda_i\xi_i + \delta_i$, $E(\delta_i) = 0$, $COV(\xi_i, \delta_i) = 0$, or without random error. So in either situation, researchers can use our figures to list the vanishing tetrads.

Independent Vanishing Tetrads

After the researcher identifies the vanishing tetrads implied by a model, the next step is to eliminate redundancy among them before he or she can conduct a simultaneous test on whether the implied vanishing tetrads are consistent with the sample data. In a simple model with four effect indicators, such as the one in Figure 1a, where all three tetrads vanish, only two of them are independent of each other. This is clearly shown in Equation 4, where any two of the vanishing tetrads imply the third; therefore, one of the vanishing tetrads is redundant and should be excluded from the test. For models with more than four variables, detection of redundant vanishing tetrads requires careful algebraic derivation.

Bollen and Ting (1993) demonstrated a general rule that whenever the same pair of covariances appears in two vanishing tetrads, a redundant vanishing tetrad is implied. Through algebraic substitution, the common terms are eliminated, and a redundant third vanishing tetrad is implied. Suppose we have

$$\tau_{1342} = \sigma_{13}\sigma_{42} - \sigma_{14}\sigma_{32} = 0, \quad (9)$$

and

$$\tau_{1452} = \sigma_{14}\sigma_{52} - \sigma_{15}\sigma_{42} = 0, \quad (10)$$

[3] Because vanishing tetrads implied by a model with correlated errors are a subset of those from the same model with no correlated errors, the two models are nested in terms of vanishing tetrads. Researchers can test whether the error terms are correlated by comparing the CTA test statistics between the two models.

with two common covariances, σ_{14} and σ_{42}, in both vanishing tetrads. Algebraic substitution leads to the redundant vanishing tetrad

$$\tau_{1352} = \sigma_{14}\sigma_{52} - \sigma_{15}\sigma_{32} = 0. \qquad (11)$$

In a model with five effect indicators, as shown in Figure 1c, only five out of the 15 implied vanishing tetrads are independent. For example, τ_{1234}, τ_{1243}, τ_{1235}, τ_{1532}, and τ_{1345} together form a set of nonredundant vanishing tetrads in this model.

Depending on the order of redundancy elimination, one could come up with different sets of independent vanishing tetrads. To ensure that the test result is not affected by the selection, Bollen and Ting (1993) suggested several sensitivity checks.[4]

Statistical Test

Once a researcher has developed the independent vanishing tetrads, it is time to test whether they are zero. The sample counterparts of the population vanishing tetrads, t_{ghij}, typically have nonzero values because of sampling errors. We need to test between the hypotheses that H_o: $\boldsymbol{\tau} = \mathbf{0}$, and H_a: $\boldsymbol{\tau} \neq \mathbf{0}$, on the basis of the sample data. Bollen (1990) derived a test statistic, T, which tests multiple vanishing tetrads simultaneously (also see Bollen & Ting, 1993).[5] The test is constructed as follows:

$$T = N\, t'\, \Sigma_{tt}^{-1}\, t, \qquad (12)$$

where N is the sample size, $\mathbf{t}$ is a vector of the independent sample tetrad differences, and Σ_{tt}^{-1} is the inverse of the covariance matrix of the limiting distribution of t as N goes to infinity. The T statistic asymptotically approximates a chi-square variate with degrees of freedom equal to the number of vanishing tetrads considered in the test. A nonsignificant result suggests that the observed tetrad differences are not significantly different from zero, indicating that the data are consistent with the vanishing tetrads implied by the hypothesized model.

The covariance matrix, Σ_{tt}, in Equation 12 can be obtained through the following steps. First, form a vector of tetrad equations, $\boldsymbol{\tau}$, which includes all independent vanishing tetrads in the test. The elements in $\boldsymbol{\tau}$ take the form of $\sigma_{gh}\sigma_{ij} - \sigma_{gi}\sigma_{hj}$. Second, build a vector, $\boldsymbol{\sigma}$, of nonredundant covariances that appear in $\boldsymbol{\tau}$. Third, construct a covariance matrix, Σ_{ss}, of the limiting distribution of the sample covariances corresponding to the elements in $\boldsymbol{\sigma}$. The elements in Σ_{ss} are given by

$$[\Sigma_{ss}]_{ef,gh} = \sigma_{efgh} - \sigma_{ef}\sigma_{gh}, \qquad (13)$$

where σ_{efgh} is the fourth-order moment for the e, f, g, and h variables. Its sample estimator is

$$S_{efgh} = (1/N)[\Sigma(X_e - \overline{X}_e)(X_f - \overline{X}_f) \cdot (X_g - \overline{X}_g)(X_h - \overline{X}_h)]. \qquad (14)$$

If the observed variables are multinormally distributed, the elements in Σ_{ss} simplify to

$$[\Sigma_{ss}]_{ef,gh} = \sigma_{eg}\sigma_{fh} - \sigma_{eh}\sigma_{fg}. \qquad (15)$$

Finally, the covariance matrix, Σ_{tt}, is estimated by

$$\Sigma_{tt} = (\partial\boldsymbol{\tau}/\partial\boldsymbol{\sigma})'\, \Sigma_{ss}\, (\partial\boldsymbol{\tau}/\partial\boldsymbol{\sigma}), \qquad (16)$$

where $(\partial\boldsymbol{\tau}/\partial\boldsymbol{\sigma})$ is the partial derivative of the vector $\boldsymbol{\tau}$ with respect to the vector $\boldsymbol{\sigma}$. Bollen (1990) also derived a modification of this test to the tetrad differences of correlation coefficients rather than covariances (also see Bollen & Ting, 1993).

In some situations there may be doubt as to whether the test statistic follows the asymptotic chi-square distribution. Bollen and Ting (1998) suggested a procedure to provide a bootstrap estimate of the p value for a given test statistic. Although full justification of the bootstrap also relies on asymptotic theory, in several areas it appears that the bootstrap approaches its asymptotic properties sooner than do the usual test statistic methods. Though this remains an area for research, the evidence from Bollen and Ting (1998) suggests that this is true for bootstrapping the tetrad test statistic. Therefore, in our small to moderate sample examples below, we report the bootstrapped-based p value in addition to the more traditional one.

When two models are nested in terms of vanishing tetrads, that is, vanishing tetrads of one model are a

---

[4] First, verify the test result with different sets of independent vanishing tetrads, and use Bonferroni correction to adjust the alpha level for multiple testing. Second, perform a test on the excluded redundant vanishing tetrads after eliminating redundancy in this group. Third, conduct tests on each redundant vanishing tetrad with Bonferroni correction for multiple testing.

[5] Spearman and Holzinger (1924), Kelley (1928), Wishart (1928), and Kenny (1974) have proposed significance tests for a single vanishing tetrad. All these tests are asymptotic, assume a multivariate normal distribution among the observed variables, and are not simultaneous tests for multiple vanishing tetrads.

subset of those in another model, these two models can be tested against each other. The more restricted model, M, implies a greater number of vanishing tetrads than the less restricted one, L. We refer to their test statistics as T_M and T_L with degrees of freedom df_M and df_L, respectively. If the two test statistics are not significantly different from each other, we prefer the model with the most vanishing tetrads; otherwise, the model with the fewest vanishing tetrads will be selected. The significance test for two nested models, T_D, is

$$T_D = T_M - T_L, \tag{17}$$

with degrees of freedom equal to the $df_M - df_L$.

Examples

Simulation Example

We start with a simulation example to examine whether the tetrad test performs as we expect when the true model that underlies the data is known. We generated a sample of 1,000 cases according to the causal indicator model in Figure 1d. The equation for this model is

$$\eta = 0.5y_1 + 0.8y_2 + 0.3y_3 + 0.9y_4 + 0.7y_5 + \zeta. \tag{18}$$

The five causal indicators are generated from normal distributions, and the covariances among them are randomly assigned.[6]

We want to compare two models. One has all five measures as effect indicators (see Figure 1c), and the other is the true model with all five as causal indicators (see Figure 1d). The latter model implies no vanishing tetrads, whereas all tetrads in the one-factor model vanish. Thus the two models are nested in terms of vanishing tetrads. The difference between the two models lies on whether $\tau = 0$ for all tetrads. In other words, testing between the two models is the same as testing the vanishing tetrads of the effect indicator model. We illustrate the steps for this test with the following CTA-SAS program (Ting, 1995):

```
%include 'c:\cta-sas2.mac';
%cta(cmatrix =
  2.034
  0.113 1.281
  0.510 0.093 1.572
  0.105 0.857 0.447 1.708
  0.998 0.228 0.170 0.345 1.651,
  n = 1000,
  vars = y1 y2 y3 y4 y5);
```

The first line of the program tells SAS to locate the file with the CTA-SAS program codes. The %cta command on the second line calls the CTA test routine, and we supply the routine with the covariance matrix, sample size, and variable list. The program will then list all tetrads, identify a set of independent vanishing tetrads, and calculate the test statistic. The user can modify the program to get the vanishing tetrads for a subset of the variables.

With five variables, there are a total of 15 tetrads. In Table 1, CTA-SAS lists all tetrads with identification numbers, tetrad values, and t-test statistics for every single tetrad. The program automatically conducts a check to identify all independent vanishing tetrads. In this example, we have five independent vanishing tetrads as listed in Table 2. Finally, the program computes a simultaneous test on these five tetrads and gives the vanishing tetrad test statistic of 64.13 with 5 df, and the P value is less than .001. In this case, a significant test statistic would lead to the rejection of the effect indicator model in favor of the causal indicator one. We alter the parameter values and generate additional models with the same causal indicator structure and get essentially the same results. In additiôn, we generate data according to the effect indicator model in Figure 1c and apply the tetrad test. Here, as expected, we find a nonsignificant test statistic of 3.88 with 5 df. The p value is .57, which is consistent with the model that generated the data. Thus, the simulated examples illustrate that the tetrad test performs as predicted.

Life Satisfaction Example

This example has the same competing models as the previous simulated data (see Figure 1, c and d) but uses real empirical data that comes from the General Social Survey (GSS; Davis & Smith, 1991). The GSS consists of a probability sample of English speaking adults in the continental United States. The 1991 survey has several indicators of life satisfaction for different areas of life. Specifically the interviewer asks the respondent to reveal his or her degree of satisfac-

[6] Each of the observed indicators is generated by $y_i = \Sigma a_{ij} y_{ij}$, where $i,j = 1 \ldots 5$, a_{ij} is a constant randomly generated from a uniform distribution, $U(-1,1)$, and y_{ij} is a random variable that comes from a normal distribution, $N(0,1)$. The covariance between y_i and y_j equals the expected value of $a_{ij}^2 y_{ij}^2$, and $a_{ii} y_{ii}$ is the unique part of y_i that is not correlated with other indicators.

Table 1
CTA-SAS Output: List of all Tetrads, Tetrad Values, and t Values

ID	Tetrad	Tetrad value	t
1	$t(1,2,3,4)$	−0.38656	−6.5116
2	$t(1,2,4,3)$	0.04075	2.0206
3	$t(1,3,4,2)$	0.42731	7.6457
4	$t(1,2,3,5)$	−0.09707	−4.2012
5	$t(1,2,5,3)$	−0.07360	−1.7135
6	$t(1,3,5,2)$	0.02347	0.4765
7	$t(1,2,4,5)$	0.01505	1.1543
8	$t(1,2,5,4)$	−0.81630	−10.6463
9	$t(1,4,5,2)$	−0.83135	−10.9048
10	$t(1,3,4,5)$	0.15810	5.3569
11	$t(1,3,5,4)$	−0.27016	−4.2543
12	$t(1,4,5,3)$	−0.42826	−7.3344
13	$t(2,3,4,5)$	−0.11361	−2.6156
14	$t(2,3,5,4)$	−0.06983	−3.2350
15	$t(2,4,5,3)$	0.04377	0.9458

Note. On the basis of five simulated variables, with $n = 1{,}000$, the confirmatory tetrad analysis (CTA) SAS program assigns each tetrad a unique identification number (ID), prints all possible tetrads (Tetrad), computes their tetrad values, and provides t-test statistics against the null hypothesis that the tetrad values equal zero.

tion with (a) place of residence, (b) hobby, (c) family, (d) friends, and (e) health. The sample covariances among these five variables are

$$S = \begin{bmatrix} 2.082 & & & & \\ 0.712 & 2.327 & & & \\ 0.223 & 0.309 & 0.324 & & \\ 0.239 & 0.366 & 0.159 & 0.296 & \\ 0.428 & 0.693 & 0.277 & 0.286 & 2.232 \end{bmatrix}. \tag{19}$$

One possibility is that the respondent has a general level of satisfaction with his or her life. The responses to the specific questions on the five areas of life will each reflect this general life satisfaction construct. Alternatively, the variables may be causal indicators that collectively determine overall life satisfaction. Either model is a substantively plausible structure.

We first test an effect indicator model with the likelihood ratio (LR) approach, which shows a chi-square value of 16.3 with 5 *df*. The p value is less than .01, suggesting a poor fit of the model. The vanishing tetrad test also indicates a poor model fit, which has a test statistic of 14.8 with 5 *df*. The tetrad test is a test of the fit of the effect indicator model that implies more vanishing tetrads than does the causal indicator model. With $p < .05$, it lends support to the causal indicator specification for the life satisfaction variables.

One possible problem is that the large sample size ($N = 1005$) for these data may lead to excessive power for the vanishing tetrad test. Because we are testing the goodness-of-fit statistic, a large sample increases the power of differentiating even a trivial amount of discrepancy and consequently rejects the null hypothesis of good model fit. For this reason, excessive power due to large sample size may lead to the rejection of the effect indicator model. We use a method proposed in Bollen and Ting (1993) to estimate the power of the statistical tests. The power is tested with respect to an alternative model that is a one-factor model with correlated measurement errors of satisfaction between family and friends. We use the likelihood estimates obtained from the factor model without correlated errors as the parameter values for the alternative models and then add a correlation from .10 to .20 between the error terms to derive the implied covariance matrix of the alternative model (Saris & Stronkhorst, 1984). Even with a correlation between the error terms of .20, the power is only .36. The same is true for the LR test, which has a power of .42. The low power suggests that the rejection of the effect indicator model is unlikely due to excessive power with respect to this correlated error model. These results reinforce the conclusion that a causal indicator model is more consistent with the data than the effect indicator model for the life satisfaction indicators.

Civil Liberty and Tolerance Example

The next set of indicators also comes from the 1991 GSS (Davis & Smith, 1991). The questions ask

Table 2
CTA-SAS Output: List of Independent Tetrads

ID	Tetrad	Tetrad value
1	$t(1,2,3,4)$	−0.38656
2	$t(1,2,4,3)$	0.04075
4	$t(1,2,3,5)$	−0.09707
6	$t(1,3,5,2)$	0.02347
10	$t(1,3,4,5)$	0.15810

Note. The confirmatory tetrad analysis (CTA) SAS program identifies five independent vanishing tetrads from a one-factor model with five effect indicators. The effect indicator model implies all 15 tetrads should be zero. Some vanishing tetrads are redundant, as they can be derived by other vanishing tetrads. Therefore, appropriate tetrad tests should be conducted among those that are independent of each other. ID = identification number.

whether a person who holds minority viewpoints should be allowed to speak in a community, teach in a college, or to have a book in the public library. The same set of questions are asked on attitudes toward someone who is (a) against religion, (b) a communist, (c) a homosexual, (d) a racist, or (e) a person who supports a military takeover of the country. We combine the scores on whether the person be allowed to perform the three above activities to measure a respondent's degree of tolerance toward each type of minority viewpoints. This leads to five indicators of tolerance, and the covariances among the indicators are

$$S = \begin{bmatrix} 1.284 & & & & \\ 0.836 & 1.324 & & & \\ 0.885 & 0.715 & 1.338 & & \\ 0.877 & 0.844 & 0.877 & 1.442 & \\ 0.742 & 0.573 & 0.771 & 0.752 & 1.287 \end{bmatrix}. \quad (20)$$

One possible model is that a respondent has an overall attitude of tolerance that determines his or her answers to all the questions. Another possibility is a causal indicator model where these five responses plus a disturbance determine the construct of tolerance. Here, too, there is not sufficient prior knowledge to distinguish which is the appropriate model. These two possibilities correspond to the models in Figure 1, c and d, respectively, and they are nested models in terms of vanishing tetrads. The tetrad test statistic for models in Figure 1, c versus d is 33.95 with 5 *df.* With the p value less than .001, the test lends support to the causal indicator specification.

We examine the possibility of excessive power due to large sample size ($N = 817$), which may lead to the significant test statistic. To assess power we formulate an alternative model that tolerance is a two-dimensional construct with effect indicators. The first dimension, measured by the first three indicators, is tolerance of people on the left, and the second dimension, measured by the final two indicators, is tolerance of people on the right. We take the maximum likelihood estimates as the population parameters for the two-dimensional model. If this alternative model is the true one, then the power of the vanishing tetrad test is .65. Using the same alternative model, the LR test has a similar level of power at .72. The moderate power, with regard to this alternative model, suggests that excessive power may be an issue for the significant test statistic.[7]

To further illustrate the procedure, we tested the two-dimensional model with effect indicators versus the two-dimensional model with causal indicators. The effect indicator model implies four independent vanishing tetrads, whereas the causal indicator model implies none. The tetrad test statistic is 24.3 with 4 *df,* $p < .001$ (the LR test statistic is 31.5 with 4 *df,* $p < .001$). Thus, the evidence still favors the causal indicator specification. Note that we cannot use the tetrad test to distinguish whether the causal indicators determine a single dimension versus two dimension constructs, for both models imply no vanishing tetrads. Tetrad equivalent models, like the equivalent models in SEMs, may not be distinguished from the data alone.

Stress Exposure Example

A line of social science research examines the latent variable of stress exposure as indicated by the experience of a variety of life events. Using data from the 1991 GSS (Davis & Smith, 1991), we look at whether individuals have experienced stressful life events related to their (a) health, (b) work, (c) finances, or (d) family. Our sample consists of 381 respondents who had experienced at least one stressful event in the past year. Table 3 lists the events under each topic. We form four indicators of stress exposure by counting the number of events experienced under each topic and obtain the following covariance matrix among these indicators:

$$S = \begin{bmatrix} 0.325 & & & \\ -0.016 & 0.583 & & \\ -0.094 & 0.016 & 1.003 & \\ 0.010 & -0.061 & -0.035 & 0.586 \end{bmatrix}. \quad (21)$$

We expect that exposure to stressful events increases the general level of stress; thus, a causal indicator

7 We remind the reader that the power of the test statistic will depend on the alternative model against which the power is being assessed. Different alternative models can lead to different power estimates. To illustrate this, we specified a model with correlated errors between the error terms for the third and fifth variables. For this model, the power is not high (.35) if the correlation is .10, but is substantial if the errors correlate at .20 (.93). These results suggest that excessive power could be an issue for a different alternative model.

Table 3
List of Survey Questions on Stressful Events

Type of event	Indicator
Health related	1. Underwent counseling for mental or emotional problems
	2. Infertile or unable to have a baby
	3. Had a drinking problem (e.g., frequently drunk, suffered from alcoholism)
	4. Used illegal drugs (e.g., marijuana, cocaine, pills)
Work related	1. Fired or permanently laid off
	2. Demoted or switched to a less favorable position
	3. Passed over for promotion
	4. Serious trouble with boss
Financial related	1. Went bankrupt (declared personal bankruptcy)
	2. Pawned or sold off valuables to make ends meet
	3. Pressured to pay bills by stores, creditors, or bill collectors
	4. Major worsening of financial condition
Family related	1. Have serious trouble with husband/wife/partner
	2. Separated from husband/wife/partner
	3. Have serious trouble with a child

Note. Survey questions on stressful events, asked in the General Social Surveys (see Davis & Smith, 1991), were grouped to reflect four different domains: health, work, finance, and family.

model would be more appropriate. Assuming that a common stress latent variable influences these events is a questionable specification, but it is implicit when such items are factor analyzed. A vanishing tetrad test is set up to test the causal indicator model versus the effect indicator one. If our hypothesis is correct, the test statistic should be statistically significant. A tetrad test shows a chi-square value of 2.9 with 2 *df* and a p value of .24. The results are contrary to our expectation. With a moderate size sample, the p value may not be accurate. As a sensitivity check we applied the bootstrap method to generate a p value of .12 (Bollen & Ting, 1998). The results are essentially the same, and we proceed with two additional checks.

First, we examine the covariance matrix to check whether near-zero covariances between indicators lead to vanishing tetrads. This is indeed the case; four out of the six correlations among the four indicators are not significant. This is consistent with causal indicators but not with effect indicators. Second, we estimate the effect indicator model. The LR test for the effect indicator model has a chi-square value of 3.8 with 2 *df* and a p value of .15. Although an LR test suggests a good overall model fit, all of the factor loadings and the variance of the latent variable are statistically insignificant ($p > .05$). These results are inconsistent with the effect indicator model but are consistent with the causal indicator specification. This example illustrates that when a tetrad test rejects a causal indicator model, one should conduct the above checks to ensure that the tetrads hold not because of the trivial subtraction between zero covariances.

Political Democracy and Industrialization Example

The last example illustrates testing for causal versus effect indicators in a simple general SEM that includes both a latent variable model and a measurement model. In addition, it illustrates how the tests can proceed even if a construct has fewer than four indicators. The covariance matrix is taken from Bollen (1989, p. 334). The path diagram is in Figure 4 and is a simplification of that in Bollen (1989, p. 324). We have three indicators of industrialization in 1960 (ξ): *ln* GNP per capita (x_1), *ln* inanimate energy consumption per capita (x_2), and percent of labor force in industry (x_3). We also have four indicators of political democracy in 1960 (η): freedom of press (y_1), freedom of group opposition (y_2), fairness of elections (y_3), and effectiveness of elected legislature (y_4). Suppose that we wish to test whether the three indicators of industrialization are causal or effect indicators. Figure 4, a and b, diagram these possibilities. With 75 developing countries, the covariance matrix for y_1 to y_4 and x_1 to x_3 is

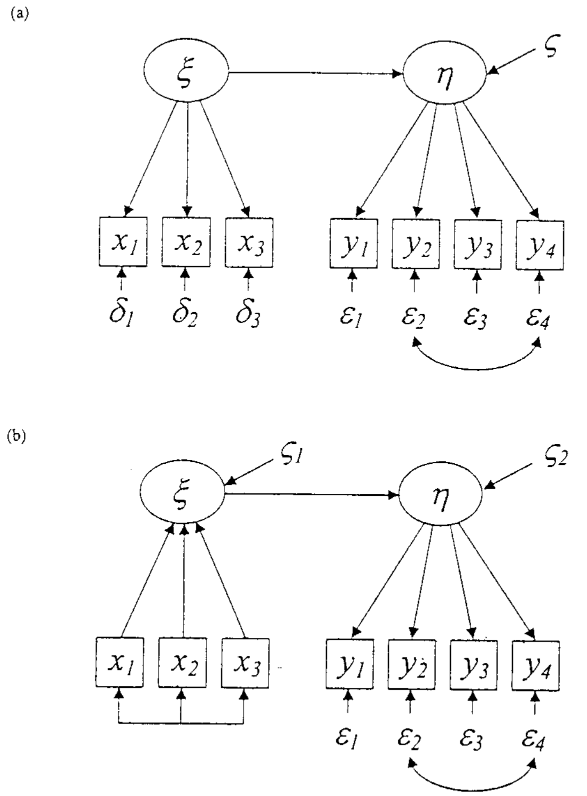

Figure 4. Industrialization and political democracy example. Models (a) and (b) are two competing conceptualizations on the measurement of industrialization, ξ. Model (a) considers x_1, x_2, and x_3 effect indicators, whereas Model (b) treats them as causal indicators. The two models are nested in terms of vanishing tetrads and can be differentiated by tetrad tests.

$$S = \begin{bmatrix} 0.537 \\ 0.990 & 2.282 \\ 0.823 & 1.806 & 1.976 \\ 0.734 & 1.273 & 0.911 & 6.878 \\ 0.619 & 1.491 & 1.169 & 6.251 & 15.579 \\ 0.786 & 1.551 & 1.039 & 5.838 & 5.838 & 10.764 \\ 1.150 & 2.240 & 1.838 & 6.088 & 9.508 & 6.687 & 11.218 \end{bmatrix}. \tag{22}$$

The tetrad test of the model with effect indicators has a chi-square test statistic of 8.82 with 12 *df*, $p = .72$, and the test statistic for the causal indicator version is 8.56 with 10 degrees of freedom, $p = 0.57$. The difference in the test statistics and degrees of freedom provides a test of the fit of the effect indicator model versus the causal indicator one. The chi-square difference of .26 with 2 *df*, $p = 0.88$, suggests evidence that the more restrictive effect indicator specification fits essentially as well as the less restrictive causal indicator one. Thus, effect indicators are a plausible structure. However, the modest sample size may affect the validity of the test. To check the sensitivity of the results, we use the bootstrap method of calculating the p value for each test statistic. The bootstrap p values for the effect indicator model and the causal indicator model are .23 and .07, respectively. The test statistic on the difference between the two models has a p value of .86. Bootstrap results yield the same conclusion as before. For precaution, we proceeded with a check on the covariance matrix to make sure that the vanishing tetrads hold not because of zero covariances. To further examine this comparison, we estimate the SEMs in Figure 4, a and b. We estimate the model with causal indicators using procedures described in Bollen and Davis (1993). The parameter estimates and standard errors from the effect indicator solution look reasonable, whereas the estimates from the causal indicator model have nonsignificant effects from the causal indicators to the latent variable of industrialization. This is further evidence in support of the effect indicator specification.

Conclusions

Much of the measurement theory in the social sciences implicitly assumes effect indicators. Though this is unlikely to change in the immediate future, there is a growing awareness of the possibility that some indicators may be determinants of, rather than reflections of, latent variables or factors. Those researchers wishing to compare a causal indicators specification versus an effect indicators specification have not had formal test procedures for this purpose. In this article we have proposed a CTA test. Frequently, the vanishing tetrads implied by a causal indicator model are nested in the vanishing tetrads implied by an effect indicator model for the same variables. As such we can compare the fit of the models with the data as well as with each other. Because the causal versus effect indicators models are not nested in the traditional sense, the usual LR tests are not applicable. Probably the most common situation will be whether to consider observed measures as all effect indicators or as all causal indicators. We present a series of typical models that researchers can use to determine the vanishing tetrads for four or more observed variables. We also provide the vanishing tetrads for the less common case of mixtures of causal and effect indicators. Furthermore, we present the vanishing tetrads for models with fewer than four indicators per latent variable and models with correlated errors of measurement. The simulation examples behaved as expected, and we illustrated the test results for several empirical examples.

Despite the promising results, we wish to close with several cautionary notes. First, we took pains to illustrate that tetrad equivalent models exist and that the tetrad test (and sometimes the LR test) cannot distinguish between such models. Second, the test statistic asymptotically follows a chi-square distribution under the null hypothesis (Bollen, 1990). Results from Bollen and Ting (1998) find that the test statistic is adversely affected by having a small N combined with a model with many parameters. The bootstrap p value for the test statistic proposed in Bollen and Ting (1998) behaves somewhat more accurately, but neither alternative is desirable when estimating large models in small samples. See Bollen and Ting (1998) for the performance of the test statistics under varying model sizes and samples sizes. Third, the quality of the test results depends on the quality of the conception of the plausible alternatives. The test, like other tests in SEMs, performs best when considerable thinking and work goes into the model before testing. We recommend that several plausible models of the relation between the latent variables and indicators be proposed in advance. Then, to the extent that their vanishing tetrads are nested, a researcher can compare their fit. As we have repeatedly argued in this article, we are not proposing a data exploration procedure that determines whether a model has causal or effect indicators. Rather we are proposing a test to help distinguish between substantively formulated alternatives that differ in their treatment of the indicators.

It is not unusual for researchers to modify their original model formulations once they have tested the model and determined that its fit is not adequate. If this is done, then the researcher moves more toward exploratory than confirmatory techniques, and we need to interpret the p values and test statistics with a great deal of caution. A researcher interested in a more exploratory approach to generating models should consider the work by the TETRAD project developed at Carnegie Mellon University. It represents one of the most ambitious attempts to search causal models based on vanishing tetrads found in statistical data (Glymour et al., 1987; Scheines et al., 1994; Spirtes et al., 1993). Extensive work has been done on the algorithm that matches causal diagrams and tetrad constraints. This exploratory approach has the objective of generating plausible models based on heuristic rules. It neither requires researchers to fully specify models in advance nor provides a formal procedure for model testing.[8]

Another cautionary note is that if support is found for causal indicators, the researcher needs to assess the identification of the model before trying to estimate it. Some guidance for this problem exists (e.g., Bollen, 1989, pp. 312–13; Bollen & Davis, 1993; MacCallum & Browne, 1993), but the literature is sparse. Finally, we should remember that the outcome of the test does not prove the validity of the specification with the best fit; other models may exist with as good or even a superior fit. Rejection of the vanishing tetrads may be due to unexamined correlated errors of measurement or unspecified direct relationships between indicators. It also is possible that a set of indicators are causal indicators with respect to one construct but effect indicators with respect to another. For instance, indicators of a child's viewing of violent television programs, playing violent video games, and listening to music with violent themes may be causal indicators of the latent variable of exposure to media violence, but the same measures could be effect indicators of another latent variable of propensity to seek violent entertainment. This is related to the naming problem that Cliff (1983) and others have described. Keeping these qualifications in mind, the tetrad test provides a useful empirical means to inform decisions on the treatment of the relation between latent and observed variables.

[8] The Tetrad II program developed by Spirtes, Scheines, Meek, and Glymour (1994) helps to identify vanishing tetrads implied by a causal model and provides Wishart's test on each sample tetrad. It does not test for multiple sample vanishing tetrads for assessing the model fit. Instead, the Tetrad II program uses the so-called Tetrad score, a heuristic index that has no known sampling properties, to select models that have the "best" match with the tetrad constraints found in the data.

References

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

Bollen, K. A. (1984). Multiple indicators: Internal consistency or no necessary relationship? *Quality and Quantity, 18,* 377–385.

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

Bollen, K. A. (1990). Outlier screening and a distribution-free test for vanishing tetrads. *Sociological Methodology and Research, 19,* 80–92.

Bollen, K. A., & Davis, W. (1993). *Causal indicators in structural equation models.* Paper presented at the American Sociological Association Convention, Miami Beach, Florida, August 1993.

Bollen, K. A., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin, 10,* 305–314.

Bollen, K. A., & Ting, K. F. (1993). Confirmatory tetrad analysis. In P. Marsden (Ed.), *Sociological methodology 1993* (pp. 147–1750). Washington, DC: American Sociological Association.

Bollen, K. A., & Ting, K. F. (1998). Bootstrapping a test statistic for vanishing tetrads. *Sociological Methods and Research, 27,* 77–102.

Cliff, N. (1983). Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research, 18,* 115–126.

Davis, J. A., & Smith, T. W. (1991). *General social surveys 1972–1991: Cumulative codebook.* Chicago: NORC.

Frydenberg, M. (1990). The chain graph Markov property. *Scandinavian Journal of Statistics, 17,* 333–353.

Glymour, C., Scheines, R., Spirtes, P., & Kelly, K. (1987). *Discovering causal structure.* Orlando, FL: Academic Press.

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.

Hayduk, L. A. (1987). *Structural equation modeling with LISREL.* Baltimore, MD: Johns Hopkins University Press.

Holmes, T. H., & Rahe, R. H. (1967). The social readjustment rating scale. *Journal of Psychosomatic Research, 14,* 213–218.

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, 70,* 631–639.

Kelley, T. L. (1928). *Crossroads in the mind of man.* Stanford, CA: Stanford University.

Kenny, D. A. (1974). A test for vanishing tetrad: The second canonical correlation equals zero. *Social Science Research, 3,* 83–87.

Land, K. (1970). On the estimation of path coefficients for unmeasured variables from correlations among observed variables. *Social Forces, 48,* 506–511.

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

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

Luijben, T. C. (1991). Equivalent models in covariance structure analysis. *Psychometrika, 56,* 653–665.

MacCallum, R. C., & Browne, M. (1993). The use of causal indicators in covariance structure models: Some practical issues. *Psychological Bulletin, 114,* 553–541.

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.

Neuberg, S. L., West, S. G., Judice, T. N., & Thompson, M. M. (1997). On dimensionality, discriminant validity, and the role of psychometric analyses in personality theory and measurement: Reply to Kruglansli et al.'s (1997) Defense of the Need for Closure Scale. *Journal of Personality and Social Psychology, 73,* 1017–1029.

Ozer, D. J., & Reise, S. P. (1994) Personality assessment. *Annual Review of Psychology, 45,* 357–388.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1,* 385–401.

Saris, W. E., & Stronkhorst, L. H. (1984). *Causal modelling in nonexperimental research.* Amsterdam: Sociometric Research.

Scheines, R., Spirtes, P., Glymour, C., & Meek, C. (1994). *TETRAD II: Tools for causal modeling (program).* NJ: Erlbaum.

Spearman, C. (1904). General intelligence objectively determined and measured. *American Journal of Psychology, 15,* 201–293.

Spearman, C., & Holzinger, K. (1924). The sampling error in the theory of two factors. *British Journal of Psychology, 15,* 17–19.

Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation, prediction, and search: Springer-Verlag lecture notes in statistics* (Vol. 81). New York: Springer-Verlag.

Spirtes, P., Scheines, R., Meek, C., & Glymour, C. (1994). *TETRAD II: Tools for causal modeling* [User's manual]. NJ: Erlbaum.

Ting, K. F. (1995). Confirmatory tetrad analysis in SAS. *Structural Equation Modeling, 2,* 163–171.

Verma, T., & Pearl, J. (1990). Equivalence and synthesis of causal models. In P. P. B. Bonissone (Ed.), *Uncertainty in Artificial Intelligence* (Vol. 6, pp. 220–227). Cambridge, MA: Elsevier Science.

Wishart, J. (1928). Sampling errors in the theory of two factors. *British Journal of Psychology, 19,* 180–187.

(Appendix follows)

Appendix

Vanishing Tetrads for Models With Fewer Than Four Indicators per Latent Variable

Sometimes we have fewer than four indicators per latent variable. In this Appendix we describe how to use indicators from other latent variables in a model to determine the vanishing tetrads implied by having causal or effect indicators.[A1] In Figure A1, we consider a latent variable, ξ_2, with three indicators and use one indicator, x_1, from another latent variable, ξ_1, to evaluate the model implied tetrads. In most cases, the causal direction between ξ_2 and x_1 has no consequence on the vanishing tetrads implied by a model. We present both cases only when the causal status of x_1 matters.

In Figure A1, Panel a is a model of two correlated latent variables with ξ_2 measured by three effect indicators. All three vanishing tetrads are implied in this model. The subsequent three models change the effect indicators to causal indicators. Only one vanishing tetrad is implied for three indicator measurement models with one causal indicator. Measurement models with two or three causal indicators imply no vanishing tetrad. Also notice that in Figure A1, Panels a through d are nested models in terms of vanishing tetrads and can be tested against each other. Panels e through h repeat the first four models with the only change being that ξ_1 now influences ξ_2. The implications for vanishing tetrads are identical to those in Panels a through d. We reverse the causal direction between ξ_1 and ξ_2 from Panels i through n. With three effect indicators for ξ_2 in Panel i, all three vanishing tetrads are implied. Changing one of the effect indicators to a causal indicator in Panel j has no impact on the vanishing tetrads. Only one vanishing tetrad remains as we add one more causal indicator to Panel l, and no vanishing tetrad is implied when all three effect indicators are changed to causal indicators in Panel n.

Panels k and m of Figure Al are particularly interesting in that they are identical to j and l, respectively, except that x_1 is a causal indicator for ξ_1 rather than an effect indicator. The change of status for x_1 leads to differences in model-implied vanishing tetrads. The vanishing tetrad for Panel k is nested in the tetrads for j, and those in Panel n are nested in the vanishing tetrads for l. This means that we can test the direction of influence of a single indicator of a latent variable under these two particular instances.

For latent variables with two indicators, we can use two indicators from another latent variable, as shown in Figure A2, to construct the tetrads. With two correlated latent variables, each with two effect indicators, only one vanishing tetrad is implied in Figure A2, Panel a. We switch the effect indicators of ξ_2 to causal indicators in Panels b and c of Figure A2, but the same vanishing tetrad is implied. We further modify these three models, with ξ_1 causing ξ_2 in Panels g to i of Figure A2. The same result holds, meaning that it is not possible to test the causal directions of a two-indicator model if the other latent variable considered in the tetrads is measured by two effect indicators. On the other hand, we can test the relationship between ξ_2 and its indicators if ξ_1 has two causal indicators as shown in Panels f, m, and o of Figure A2.

[A1] Alternatively a researcher can use computational algorithms from Glymour et al. (1987) and Spirtes et al. (1994) to derive the vanishing tetrads implied by recursive linear SEMs. Their subsequent work further provides conditions where vanishing tetrads can be computed for nonrecursive linear SEMs. Because not all readers will have access to their programs, we provide a discussion of a variety of common models and list the implied vanishing tetrads here for easy use.

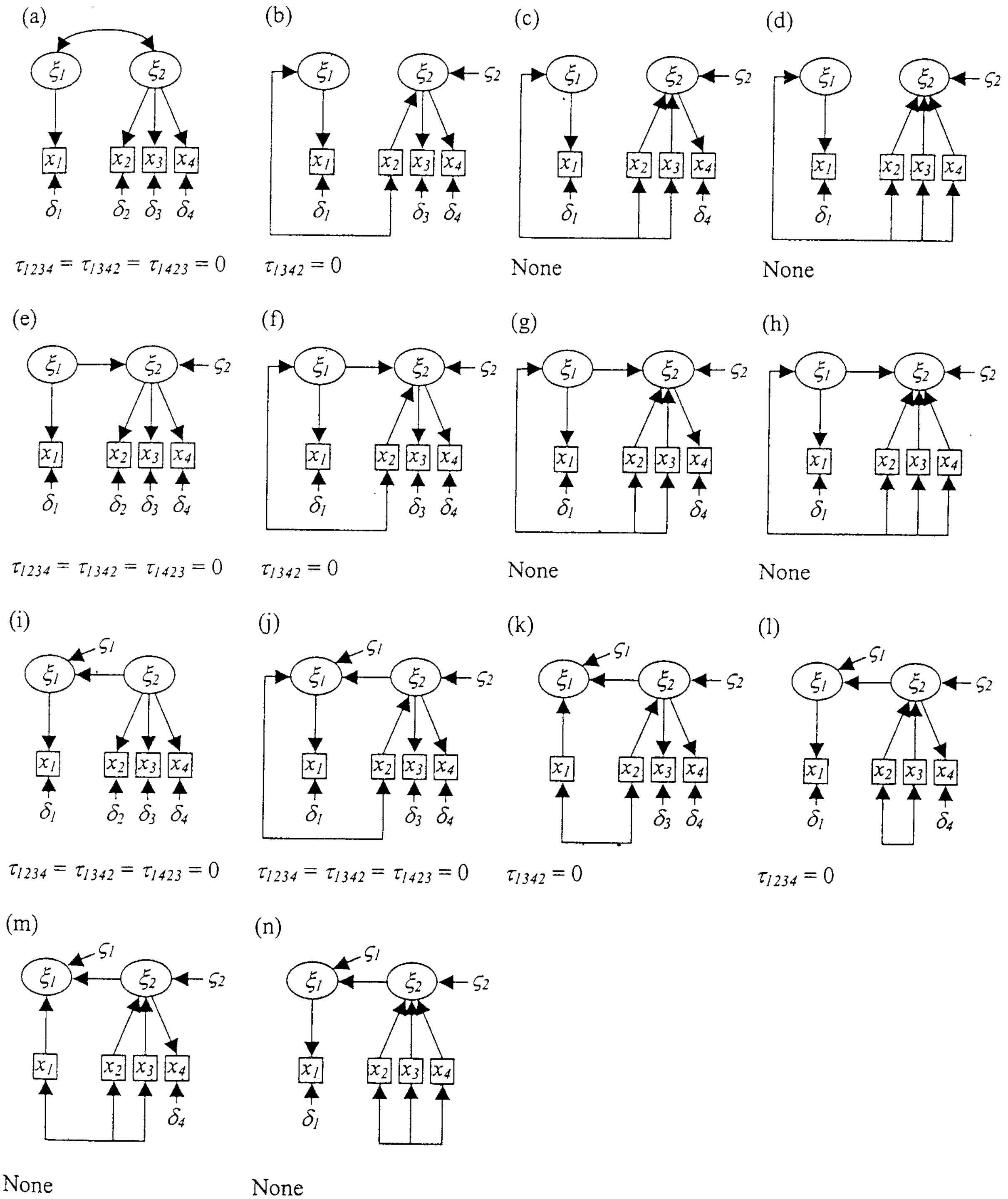

Figure A1. Illustration of how to use a single indicator from one factor in conjunction with three indicators from another for vanishing tetrad tests for causal versus effect indicators.

(*Appendix continues*)

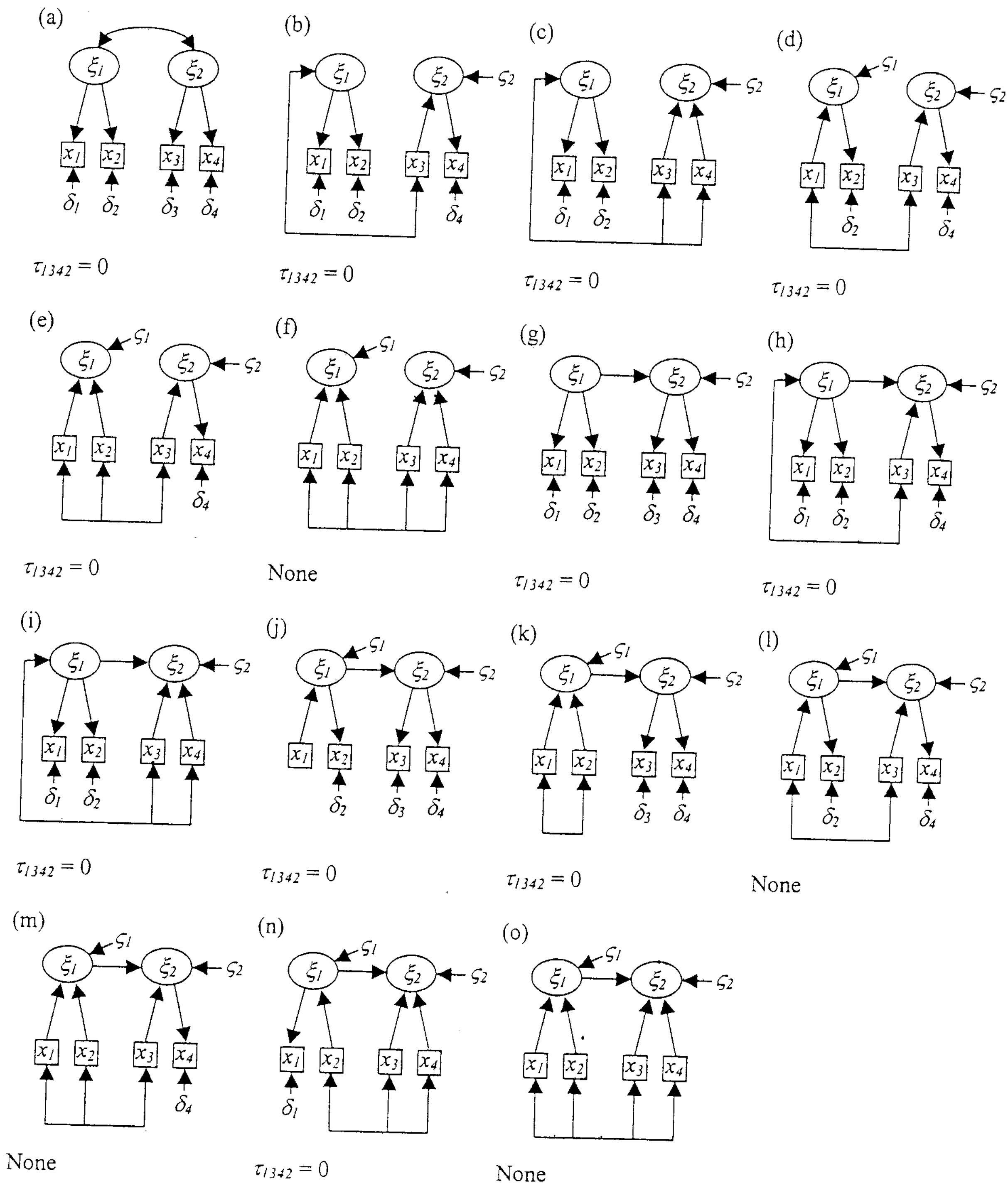

Figure A2. Illustration of the use of pairs of indicators from different factors to test for causal versus effect indicators. The implied vanishing tetrads are below the path diagrams.

Received April 13, 1998
Revision received October 22, 1999
Accepted November 4, 1999 ■