MIS Quarterly

RESEARCH NOTE

REVISITING BIAS DUE TO CONSTRUCT MISSPECIFICATION: DIFFERENT RESULTS FROM CONSIDERING COEFFICIENTS IN STANDARDIZED FORM[1]

Miguel I. Aguirre-Urreta
School of Accountancy and MIS, College of Commerce, DePaul University,
1 East Jackson Boulevard, Chicago, IL 60604 U.S.A. {maguirr6@depaul.edu}

George M. Marakas
School of Business, University of Kansas, Summerfield Hall,
1300 Sunnyside Avenue, Lawrence, KS 66045 U.S.A. {gmarakas@ku.edu}

Researchers in a number of disciplines, including Information Systems, have argued that much of past research may have incorrectly specified the relationship between latent variables and indicators as reflective when an understanding of a construct and its measures indicates that a formative specification would have been warranted. Coupled with the posited severe biasing effects of construct misspecification on structural parameters, these two assertions would lead to concluding that an important portion of our literature is largely invalid. While we do not delve into the issue of when one specification should be employed over another, our work here contends that construct misspecification, but with a particular exception, does not lead to severely biased estimates. We argue, and show through extensive simulations, that a lack of attention to the metric in which relationships are expressed is responsible for the current belief in the negative effects of misspecification.

Keywords: Construct specification, formative, reflective, simulations, standardized coefficients, unstandardized coefficients

Introduction

It would be rather uncontroversial to state that the *raison d'être* of an academic field of study, such as the Information Systems discipline, is the development and testing of theory about relationships between units of interest. When attempting to empirically test their theoretical propositions, researchers must find ways to make these constructs manifest, in order to ascertain whether the expected relationships among them hold as predicted by the proposed research model.

Conclusions about relationships between such constructs of interest, however, are subject to the adequacy of the indicators that were chosen to empirically test the theory or theories under examination. This falls under the general heading of validity. While most of the work in this area has generally assumed that constructs were modeled in a reflective manner, researchers have started discussing an alternative specification of the relationship between constructs and their manifest indi-

[1]Joseph Valacich was the accepting senior editor for this paper. Andrew Burton-Jones served as the associate editor.

The appendices for this paper are located in the "Online Supplements" section of the *MIS Quarterly*'s website (http://www.misq.org).

cators, where the latent variables are caused by the latter. Much of the emergence of this specification in the behavioral and social sciences was prompted by the work of Diamantopoulos and Winklhofer (2001), which draws on earlier work by Bollen (Bollen 1984; Bollen and Lennox 1991), Cohen et al. (1990), and Curtis and Jackson (1962), among others. Formative and reflective are thus the two currently accepted ways of specifying the relationship between latent constructs and observed variables that are causally related to them.

Researchers from a variety of backgrounds have engaged in extensions to this research and debates about the nature of formative indicators. Special issues have recently been published on this topic in both *Psychological Methods* (Volume 12, Issue 2, 2007) and *Journal of Business Research* (Volume 61, Issue 12, 2008). Within the IS field, the seminal work of Petter et al. (2007) has been extensively cited in the short time since its publication. A debate has also ensued as to the appropriate modeling of the computer self-efficacy construct (Hardin et al. 2008a, 2008b; Marakas et al. 2007, 2008), and other articles have recently appeared on the topic (Cenfetelli and Bassellier 2009; Kim et al. 2010). There is clearly interest in understanding and appropriately employing this methodological development in applied research.

Several researchers have conducted extensive reviews of the empirical literature in the fields of marketing and consumer research (Jarvis et al. 2003), organizational behavior (MacKenzie et al. 2005), leadership (Podsakoff et al. 2003), and information systems (Petter et al. 2007). In these exemplars, a set of guidelines for the appropriate use of formative and reflective specifications was developed, and then published research was analyzed to assess whether researchers had properly specified the relationship between latent variables and observed indicators. Also, in several of these articles, Monte Carlo simulations were used to illustrate the deleterious effects of construct misspecification on the estimation of structural parameters of interest. The scenario under consideration here occurs when the relationship between a latent variable and its indicators has been modeled as reflective when, unbeknownst to the researcher, the relationship is truly formative; that is, the researcher has misspecified the relationship between latent and manifest variables in a portion of the research model. Two main conclusions emerge from these studies. First, an apparently large portion of empirical studies in these disciplines had been conducted with models that included one or more misspecified constructs. Second, the effects of misspecification on the relationships among latent variables are quite significant, leading to major bias in their estimation—bias defined as an obtained result that systematically differs from its true value in the population of interest. As a result, researchers may conclude from these statements that a large portion of the empirical research in these areas, including the discipline of information systems, is largely invalid.

While the issue of when and whether constructs are misspecified, and many other conceptual and interpretational issues have received a fair amount of attention (although much remains to be explored), the conclusion that misspecification leads to severe bias in estimation has remained largely unexplored and unchallenged. Our examination of research in this area suggests that a lack of attention to the metric of latent variables is responsible for the posited bias, and when considering the relationships in their standardized form neither the direction nor the magnitude of relationships are biased to the degree previously discussed. In addition, we believe the extensions of past research conducted by Petter et al. focusing on Type I and II errors present some limitations and therefore do not accurately reflect the occurrence of these two important issues.

We conducted a replication of the work by Jarvis et al. (2003) and by Petter et al. to illustrate our arguments and highlight the quite different conclusions that can be reached when considering the standardized form of the coefficients under examination. We contend, and show with our results, that in most cases the apparent bias was a function of not carefully considering the metric in which relationships were expressed, rather than actual bias in the underlying relationship. One exception remains for the case of formatively specified constructs, and then only for a single path. While we believe the first of the two conclusions referred to above is still subject to discussion, we present evidence here that tentatively refutes the second. Research on the proper specification of constructs is still of critical importance to uncover the true relationships between variables and their structure, but the consequences of misspecification seem to be much less dire than previously thought. Our conclusions are, of course, limited to the particular models investigated, but we believe they have important implications for the use of latent variables in general.

The rest of this work is organized as follows. First, we present and develop the main thesis of our discussion, namely the importance of identification and scale setting issues in structural equation modeling research, and their implications for interpreting results from simulations like the ones conducted here. Second, the effects of construct misspecification in research models with latent variables are developed, and a complete replication and extension of past research is conducted. Finally, we discuss our findings and draw implications for researchers using these methods.

Identification in Confirmatory Factor Analysis

Confirmatory factor analysis (CFA) models, like the ones explored in this research, need to satisfy two conditions in order to be identified.[2] First, the number of free parameters must be less than the number of unique elements in the moment matrix being analyzed. That is, the degrees of freedom of the proposed model must be greater than or equal to zero. Second, every latent variable, including disturbances, factors, and measurement errors, must be assigned a scale or metric (Kline 2005). Otherwise, unless constraints are imposed, it is always possible to find different parameter matrices that will reproduce the observed covariance matrix equally well (Millsap 2001). A model is thus identified when the constraints set on the covariance matrix imply a unique set of parameter values.

In addition to the conventional practice of setting the scale of disturbances and measurement errors through unit loading constraints, there are generally two alternative approaches for obtaining latent variable identification. The first and most commonly used approach for setting the scale of a latent variable involves fixing the loading of the factor on one of its indicators to a nonzero value, most generally one (however, any other value would equally do); the remaining loading estimates will be proportionally adjusted as a result. Gerbing and Hunter (1982) discussed this issue, and noted the following:

> Fixing the value of a factor loading at an arbitrary value determines the variance of the corresponding latent variable for computational purposes, but this variance is an arbitrary and meaningless metric which is neither the variance of the observed variable nor is it the actual variance of the latent variable. Consequently, the computed covariance matrices of the latent variables and of the latent and observed variables are expressed in a metric which does not correspond to the parameters of the model used to generate the data (p. 425).

By fixing a particular loading to a nonzero value, the estimated variance of the latent variable becomes a function of the variance of the common part of the manifest variable that has the imposed constraint, where the variance of the common part itself is determined by the choice of other variables included in the measurement model for that factor (Steiger 2002). It should also be noted that, unless the nonzero value chosen for the constraint happens to match the true population value of that loading, the other estimated parameters in the model will appear to be biased when compared against their population counterparts; however, this is the result of the metric in which they are expressed, and are equivalent otherwise. This approach is most commonly used in multiple group comparisons and when setting the scale for endogenous latent variables (see below).

The second approach relies on setting the variance itself to a fixed, nonzero value. While technically this approach would be applicable to both exogenous and endogenous latent variables, most commonly used software packages only allow users to directly specify the variance of exogenous latent variables. Because the variances of endogenous latent variables are nonlinear functions of other parameters in the model, they require the introduction of complex constraints, one for each endogenous variable, to be set to a fixed value. Therefore, this approach is most commonly applied only to exogenous latent variables.

In this case, instead of relying on the choice of manifest indicators loading on each factor and their common variance, researchers can opt to set the variance of a latent variable at any given number, again most commonly one, but any nonzero number would be acceptable. If the variance of both exogenous and endogenous factors were fixed at one, the resulting regression coefficient expressing the relationship between them would be in standardized form. Given the difficulties noted above for fixing the variance of endogenous factors, this is hardly the case in practice.

To summarize the discussion up to this point, latent variable models require the specification of constraints on some parameters in order to obtain unique estimates for of all the non-constrained ones. The two common approaches to setting the scale, or metric, of latent variables include fixing a selected loading to a nonzero value and fixing the variance of the exogenous latent variable to a nonzero value. In both cases, the commonly used value is one, but this is not required and any nonzero value will result in models that fit the data equally well. Importantly, different approaches to model identification and scale setting, as well as different nonzero values used for this purpose, will result in markedly different unstandardized estimates of the relationship between latent variables, even if the alternative models will fit equally well; however, when standardized, the estimates will all be identical.

[2] Readers less familiar with identification and scale issues may want to consult Appendix A before moving further.

Standardized estimates,[3] then, are particularly useful to detect the presence of bias arising from alternative modes of estimation, if any, since they do not depend on the particular parameterization of the research model. This is an important distinction, since the posited negative effects of construct misspecification on parameter estimates have been based on bias in the unstandardized structural relationships among latent variables and, as we have discussed in this section, those estimates cannot be meaningfully interpreted absent knowledge about the involved variances. We discuss the effects of this issue on results arising from past research next.

The misspecification of the relationship between the observed and latent variables will have an important effect on the estimated variance for the involved latent variables. This point was recognized and developed in great detail by MacKenzie et al. (2005) for a simple two factor model; however, the authors did not extend their discussion of these effects to consider whether there would be any consequences for the parameters when standardized, and consequently on the variance explained in the endogenous variable.

In general, construct misspecification will lead to estimates of the variance of the focal variable that are different than those at the population level. That would also the be case, however, if the models were correctly specified, but a value other than that present at the population level were used to set the metric of the latent variable, as previously discussed. In their discussion, MacKenzie et al. indicated that when a misspecified construct is in an exogenous position[4] its variance will be reduced and, with no misspecification of any other variables in the model, result in an upward bias of the unstandardized regression parameter to compensate. The converse is true when the misspecification occurs for a latent variable in an endogenous position, resulting in downward bias in the unstandardized regression estimate.

These effects are dependent on the particular nonzero values used to set the scale for the misspecified latent variable, and therefore caution is required before making any generalizations. In the particular models examined in this article, when construct misspecification occurred, the resulting reflective latent variables had their scale set by fixing the loading of one indicator at one. As a result, the estimated variance of the misspecified latent variable was indeed much lower and, as a result, the unstandardized regression estimates differed significantly from their population counterparts, as shown by Jarvis et al. (2003) and later replicated by Petter et al. (2007). These are the outcomes that have lead researchers to conclude that construct misspecification results in misleading and largely biased estimates. If, on the other hand, the latent variable misspecified as reflective were to be identified by fixing a loading to a value much higher than one, the resulting variance would be higher than before and thus the unstandardized regression estimate would be, apparently, biased downward from the known population value. Moreover, if the misspecified latent variable was positioned as exogenous in the research model, a researcher could choose to set the scale by fixing the variance at one (or any other nonzero value), which would result in yet another, different set of unstandardized results and estimates of bias.

Notably, if by sheer chance a researcher were to choose a value for the variance that were to match the one that should have been obtained had the latent variable been properly specified as formative, no bias in the unstandardized estimates would occur, even if the model was arguably misspecified. For the sake of illustration, we show this possibility when we conduct our simulations below. While a reader could argue it would be highly unlikely that a researcher were to choose that exact same value, we would contend it is also highly unlikely that any of these values is exactly one at the population level, yet we regularly choose to fix loadings or variances at one out of convenience.

Noting that unstandardized estimates are adjusted upward or downward as the variance of the latent variable changes as a result of misspecification (MacKenzie et al. 2005), and leaving aside the effects of the different identification approaches and constraints just discussed, we wondered whether construct misspecification truly leads to bias in the relationship between latent variables (e.g., the standardized estimate and variance explained by one variable on the other), or the apparent bias was a result of unstandardized relationships being expressed in different metrics. If correct, then it would be possible for past research to have greatly overstated the bias arising as a result of misspecification. We show this to be the case through a replication and extension of the work previously published by Jarvis et al. and by Petter et al.

[3]Standardized parameters figure prominently in applied research. For instance, guidelines for construct validity emphasize the need for standardized loadings to be above a certain threshold, most generally 0.70. As well, all regression estimates obtained from applications of partial least squares (PLS) are expressed in standardized form, unless the default settings of the software packages (PLS-Graph, SmartPLS) are modified. Researchers typically report only standardized estimates in published work.

[4]From a terminological standpoint, we believe the labeling of formatively specified latent variables as exogenous is misleading, since these are in turn endogenous to their own indicators. We have, however, kept our discussion consistent with past research in this regard.

Replication and Extension with Standardized Estimates

In this first set of simulations we replicate and extend past work by Jarvis et al. (2003) and highlight how the large bias previously reported is more a function of not taking into consideration issues of scale for the latent variables than a result of bias in the underlying relationships. The same correctly and incorrectly specified models are employed, which are shown in Figure 1. These same models were used by Petter et al. (2007) in their replication and we believe are representative of the degree of complexity found in actual research.

Data generation and analysis for all models in this simulation was performed using EQS 6.1 (Bentler and Wu 1995). A total of 1,500 replications each of datasets with 250 and 500 observations were drawn from the correctly specified models shown in Figure 1 (results obtained from samples of size 250 did not substantively differ from those reported here and are included in Appendix E). All of these converged properly and were included in the analyses. The first item of interest is the percent relative bias in the estimation of the structural parameters when models are misspecified. These results, as noted before, have been the main driver behind the posited negative effects of model misspecification in empirical research. Figure 2 compares previous findings (in unstandardized form) by both Jarvis et al. and Petter et al. for Models 1A, 1B, and 1C with our replication in both unstandardized and standardized form.

The differences are notable. While results from our replication, in unstandardized form, closely resemble those previously obtained (thus validating the adequacy of our data generation and analysis process), bias in the relationships among factors all but disappears when expressed in standardized form. Thus, while the posited bias in the relationship between ξ_1 and η_1 and η_3 ranges from 485 percent to 337 percent (Current Results: Unstandardized in Figure 2), depending on the parameter and model under consideration, the same data and results, when expressed in standardized metric, show very limited bias, in the 7 percent to 2 percent range (Current Results: Standardized in Figure 2). The key to interpreting this result is to remember that relationships in unstandardized form are expressed in a metric that is dependent on the particular estimated variances for the factors involved, whereas those expressed in standardized form have been rescaled to make the variances of all latent and manifest variables equal to one.

To better clarify, we work out one example in more complete detail. At the population level in Model 1C (see Appendix B for formulas and variance decompositions), the variance of the formatively specified construct (ξ_1) is 12.900, that of η_1 is 1.661 and the regression coefficient linking both factors is 0.300. The latter should be interpreted in the context of the variances of the constructs involved, and therefore ξ_1 explains 69.90 percent of the variance in η_1. When rescaled so that the variances of both latent variables equal one, the standardized path between both variables equals 0.836, which also results in 69.90 percent of explained variance (the variance of the exogenous variable equals one as a result of the rescaling, and therefore $1 \times 0.836^2 = 0.6990 = 69.90\%$).

In our replications, the average unstandardized path between these two latent variables equals 1.310, which results in the apparent bias of 337 percent shown in Figure 2 and in past research [(1.310 – 0.300) / 0.300 × 100]. The average estimated variance for ξ_1 was 0.7062 and for η_1 1.6617. Using these average variances and average unstandardized regression coefficient, the average variance explained should amount to 72.93 percent [$(0.7062 \times 1.310^2) / 1.6617$]; the observed average variance explained was 72.68 percent, the insignificant difference due to rounding errors and working with averages over 1,500 replications. The standardized coefficient, averaged over the replications, equals 0.852, which produces the bias of 2 percent [(0.852 – 0.836) / 0.836 × 100] reported in Figure 2. Using this figure we arrive at the same value for the average variance explained in η_1, namely 72.59 percent ($1 \times 0.852^2 = 0.7259 = 72.59\%$).

Therefore, by taking either the unstandardized coefficient in the context of the estimated variances for both factors involved, or the standardized coefficient when the variances have been rescaled, the average observed explained variance in both cases is the same, which it must be, since the latter represents just a rescaling of the former. The bias due to misspecification, however, is markedly different: 337 percent when using unstandardized values versus 2 percent when using standardized ones. When considering the metric in which different figures are expressed, the reconciliation of this difference is straightforward. At the population level, as specified in Model 1C, the unstandardized coefficient of 0.300 was expressed in the context of the variance of ξ_1 being 12.990 and that of η_1 being 1.661; in the results from the simulations, the average (and apparently biased) unstandardized coefficient of 1.310 is related to the estimated variance of ξ_1 being 0.7062 and that of η_1 being 1.6617. The two coefficients are clearly expressed in different metrics and are thus not directly comparable.

The standardized coefficients for both the population model and resulting from the simulations, on the other hand, are expressed in a comparable metric, and can thus be assessed for bias due to the misspecification of ξ_1 in a reflective manner. The R^2 statistics calculated above also highlight the

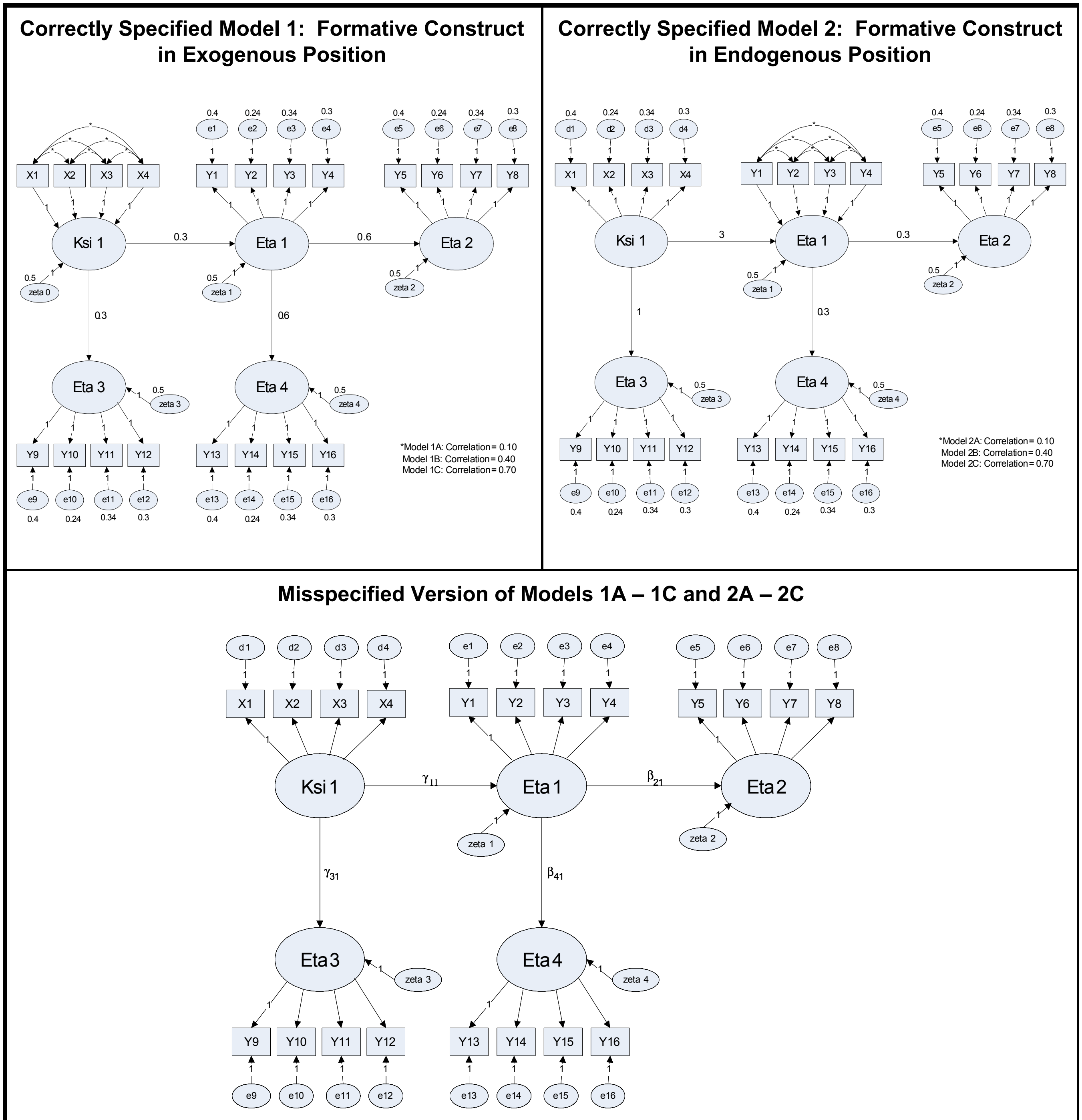


Figure 1. Models Specified in this Research (Figure 3 from "A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research," C. B. Jarvis, S. B. MacKenzie, and P. M. Podsakoff, *Journal of Consumer Research* (30), September 2003, p. 210. Copyright © 2003, University of Chicago Press. Used with permission.)

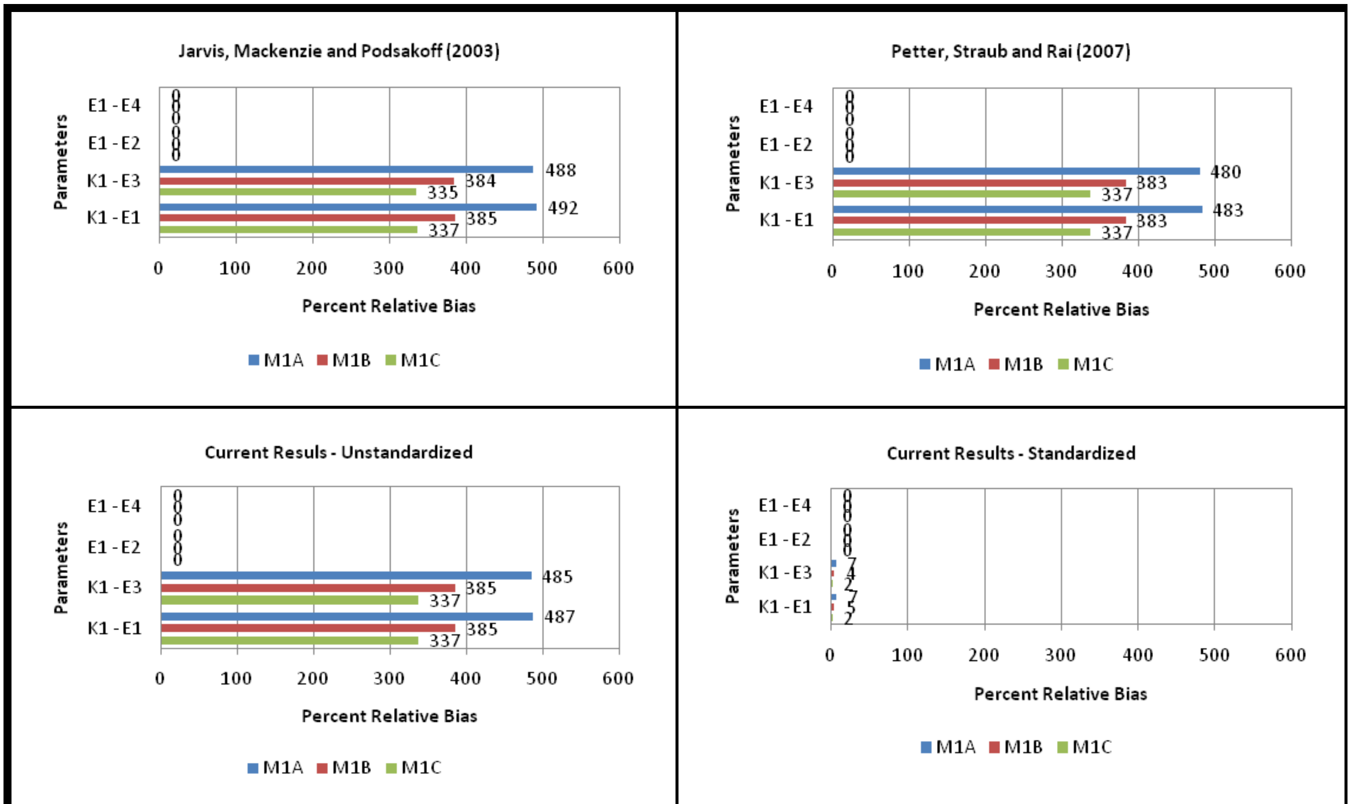


Figure 2. Percent Relative Bias in Structural Parameter Estimates for Models 1A, 1B, and 1C (exogenous formatively specified in the population)

Note: The formative item intercorrelations are 0.10 for Model 1A, 0.40 for Model 1B, and 0.70 for Model 1C.

rather problematic nature of comparing coefficients expressed in different metrics. Consider that, at the population level, the proportion of variance explained by ξ_1 in η_1 in Model 1C, as noted above, is 69.90 percent; and that the average estimated R^2 from the simulation was, for the same relationship, 72.68 percent. This small upward bias of 3.98 percent [(0.7268 – 0.6990) / 0.6990 × 100] clearly does not reconcile with the reported bias of 337 percent when using unstandardized figures. It does, however, reconcile when considering results reported in the standardized metric, since the bias corresponds to the effects of the difference between the population and observed regression coefficients when rescaled.[5]

To better illustrate the importance of considering the metric of latent variables, we also re-estimate the data generated for Model 1C by misspecifying the ξ_1 construct in a reflective manner, but set the scale by fixing its variance to 12.900, the same value that would have been obtained had the construct been properly specified as formative. While the population value for the variance would be unknown to researchers, we wanted to provide an example of how unstandardized coefficients are influenced by the variances of the latent variables and, if adjusting them, one could faithfully reproduce population unstandardized values even in the face of misspecification. Results from this exercise, reported in Table 1, show this to be the case.

The same data used to fit the misspecified version of Model 1C reported above was used in this example. Both that model and the one discussed here exhibited identical model fit (average $\chi^{2(166)}$ = 175.033 in both cases), which underscores the fact that both models are completely equivalent, and only differ in the method used to set the metric of the latent variables. As shown in Table 1, when the variance of the misspecified variable is set to the correct population value, the effects of misspecification on the unstandardized coefficients

[5]This reconciliation works as follows: the population R^2 of 0.6990 less than the average observed R^2 of 0.7268 equals 0.0278; the difference between the population standardized estimate squared $0.836^2 = 0.6989$ less the average observed standardized estimate $0.852^2 = 0.7259$, which equals 0.02701—the difference due to rounding errors in the estimates.

Table 1. Results from Misspecified Model 1C with the Variance Fixed to Population Value

Parameter	Unstandardized Population Value	Unstandardized Average Value over Replications	Percent Relative Bias
VAR (ξ_1)[a]	12.900	12.900	—
VAR (η_1)	1.661	1.662	0%
VAR (η_2)	1.098	1.098	0%
VAR (η_3)	1.661	1.662	0%
VAR (η_4)	1.098	1.097	0%
γ_{11}	0.300	0.306	2%
γ_{31}	0.300	0.306	2%
β_{21}	0.600	0.600	0%
β_{41}	0.600	0.600	0%

[a]This value was fixed to identify the exogenous factor in the replications.

disappear, save for sampling error. The standardized coefficients for this model, on the other hand, were exactly the same as for the misspecified version of Model 1C reported in Figure 2. This short example highlights how apparent bias due to misspecification is due more to its effects on the metrics of the latent variables involved than on the underlying relationship, to the extent that careful manipulation of the metric makes a misspecified model seemingly unbiased when it comes to unstandardized coefficients. Their standardized counterparts, on the other hand, are not affected by these changes.

To summarize, when the relationship between two latent variables is expressed in unstandardized form, the interpretation of its magnitude should be done taking into consideration the estimated variances of the variables involved. Otherwise, comparisons across estimates obtained using different methods of identifying the latent variables or, as is the case here, different specifications of the relationship between a latent variable and its indicators, are not meaningful and can be misleading. The alternative suggested here is to perform this comparison by expressing relationships in a common standardized metric, which results from rescaling the relationships between variables from the original metrics into those that result from setting all variances of latent and observed variables to equal one. Seen in this light, *the apparently large bias that results from the misspecification of the relationship between latent variables and indicators is more a function of comparing coefficients expressed in different metrics rather than bias in the underlying effects themselves*.

Bias arising from results expressed in standardized form is also strikingly different than previously thought when considering the misspecified versions of Models 2A, 2B, and 2C (see Figure 1), with the exception of the path linking ξ_1 to η_1. The average results for the other three paths show no noticeable bias. For the path between ξ_1 and η_3 our results are consistent with those previously published. The relationships between η_1 and η_2 and η_4 also show virtually no bias when expressed in standardized form, while we were able to replicate the apparent large bias in the estimates when those were expressed in unstandardized form, as shown in Figure 3. Although bias, even in standardized form, is not as large as was previously published, it is nonetheless very significant, and increases as the intercorrelation among the formative items increases, a finding not reported in previous research. Our results are shown in Figure 3 and contrasted with those of Jarvis et al. and Petter et al.

Although we raise some conceptual objections to the use of formatively specified constructs in a position endogenous to others in conceptual models in our discussion, it is nonetheless worth exploring the issue further. Of all paths that were reported as suffering severe bias due to misspecification of the formatively specified construct, this is the only case where bias remains even after comparisons between paths have been expressed in a common metric.

Type I and Type II Errors in Misspecified Models

In their recent article, Petter et al. (2007) sought to extend the work of Jarvis et al. (2003) by focusing on the occurrence of Type I and Type II errors in both properly and misspecified models, using the same research models previously discussed. Petter et al. concluded that "both Type I and Type II errors are

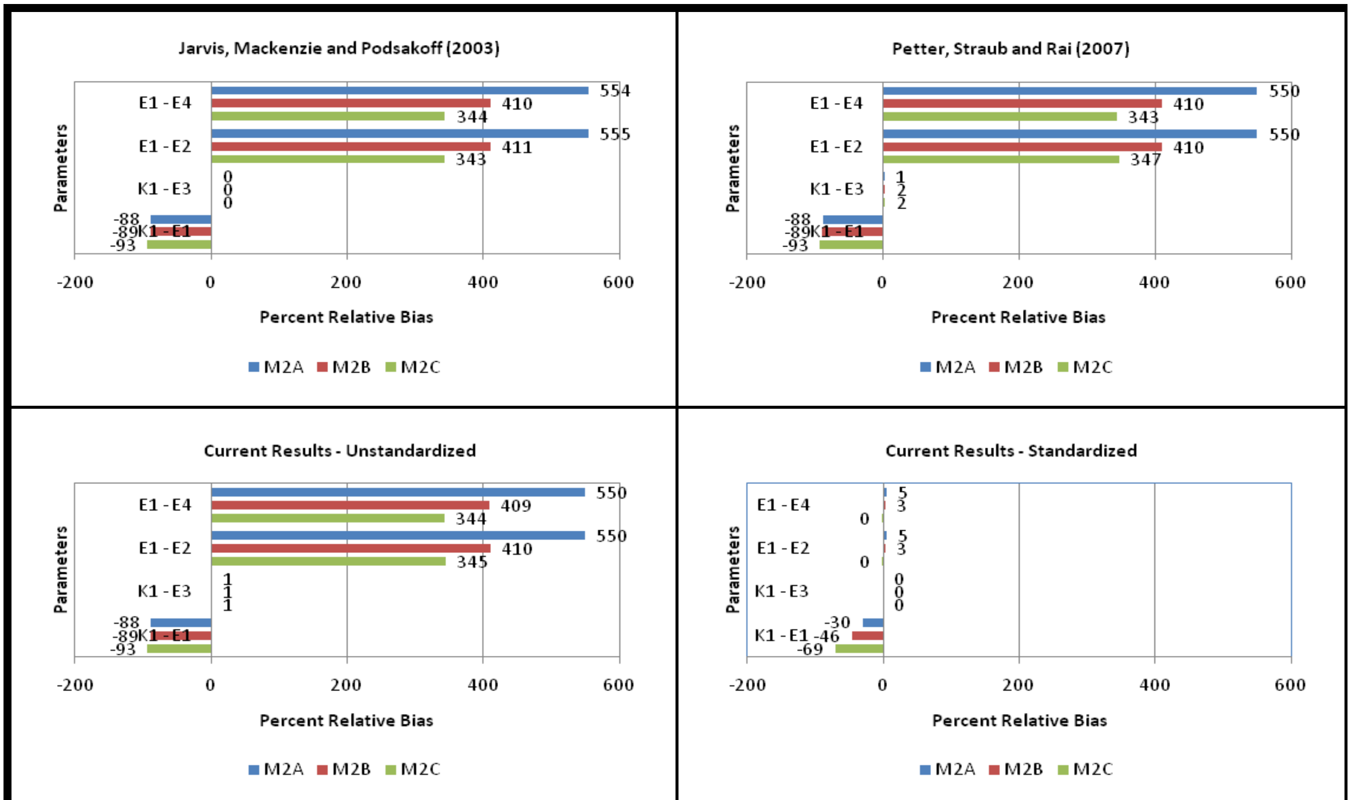


Figure 3. Percent Relative Bias in Structural Parameter Estimates for Models 2A, 2B, and 2C (endogenous formatively specified in the population)

Note: The formative item intercorrelations are 0.10 for Model 2A, 0.40 for Model 2B, and 0.70 for Model 2C.

prevalent when such misspecification occurs" (p. 644). A critical examination of their research shows our conclusions to be somewhat different than those originally published. The occurrence of the two types of errors had not been previously explored, and is clearly an important area of research. We believe, however, that limitations in the approach taken by the authors prevent fully and accurately understanding the existence and extent of these two issues.

A Type I error, or false positive, is said to occur when a significant result is obtained while the true population value of the parameter under consideration is zero (e.g., when there is no underlying relationship between the variables of interest). Type II errors, or the occurrence of false negatives, happen when a nonsignificant result is obtained while, at the population level, there is a nonzero relationship between the two variables of interest. In the context of Monte Carlo simulations, like the ones conducted in this and previous referenced research, significance (or lack thereof) is assessed at the level of each individual parameter in each individual replication, and the number of significant occurrences is then tallied. Petter et al., on the other hand, assessed Type I and II errors as properties of a set of replications as a whole by comparing the average estimate for each parameter over all their replications with the average standard error over the same, and assessing whether their ratio exceeded the critical value of a *t* statistic at the conventional alpha level of 5 percent. We contend that this focus on the aggregate level is not consistent with the nature of the concept of Type I and II errors.

This discussion also highlights a second issue with their examination of Type I error. Instead of testing a relationship that was nonexistent at the population level (e.g., a regression path equal to zero), the authors introduced a "nonsignificant" path of 0.03 between the ξ_1 and η_4 latent variables in their Models 3 (A, B, and C) and 4 (A, B, and C), shown in Figure 4. While the distinction between 0 and 0.03 might seem trivial and overtly contentious, recall from an earlier discussion that unstandardized regression coefficients cannot be interpreted in their magnitude without consideration of the variances of the involved variables. Rather, when considering this path in relation to the variances of the latent variables in-

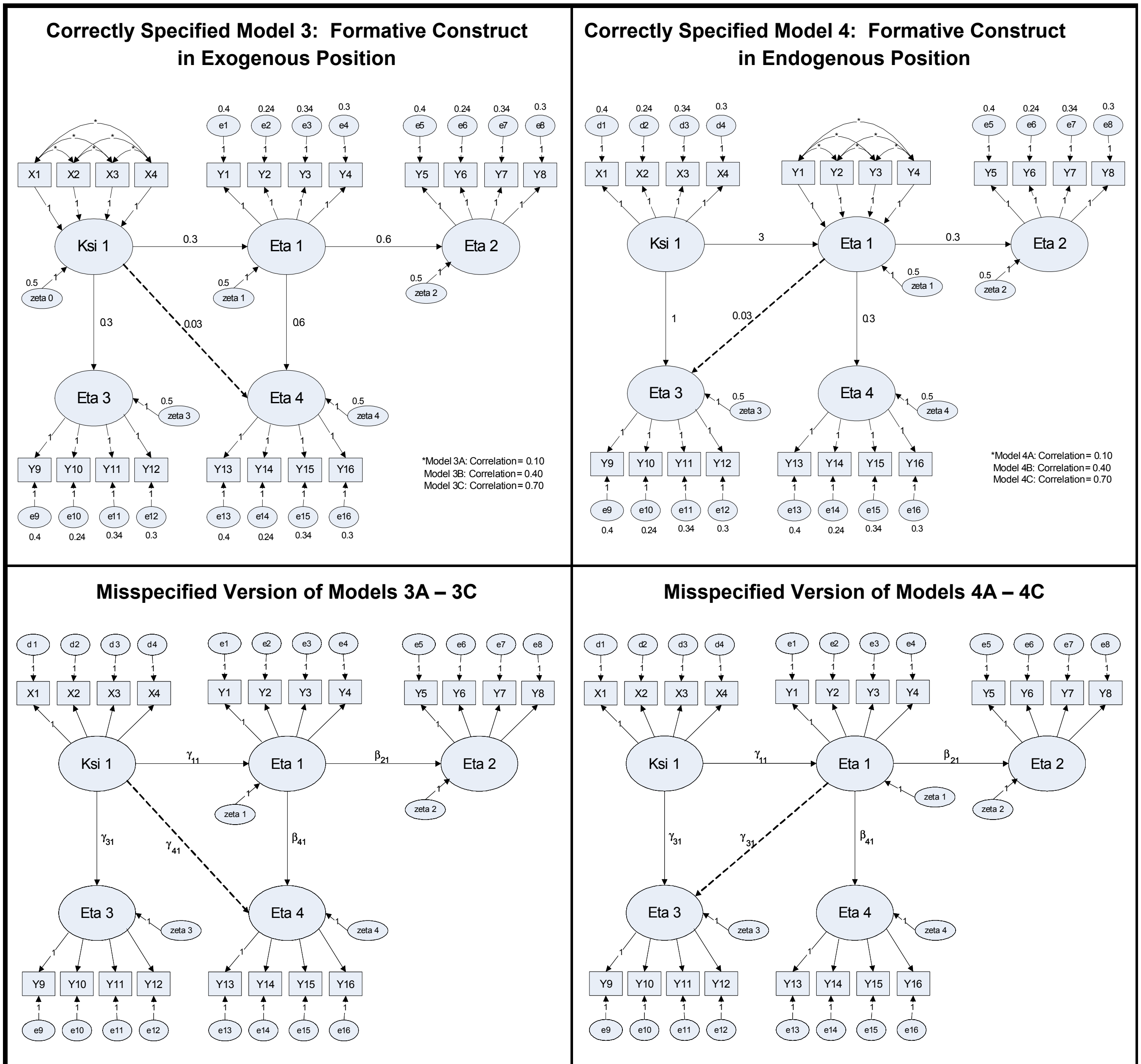


Figure 4. Models Specified by Peter et al. for their Type I Error Simulation (Figure A2 in "Specifying Formative Constructs in Information Systems Research," S. Petter, D. Straub, and A. Rai, *MIS Quarterly* (31:4), 2007, p. 651).

volved, the proportions of explained variance in the dependent variable are 0.6 percent, 0.8 percent, and 1 percent for Models 3A, 3B, and 3C, respectively, and 1 percent, 1.5 percent, and 2 percent for Models 4A, 4B, and 4C, respectively. Therefore, while the resulting effects are indeed small, they are different from zero and, since testing was done on nonzero effects, any results obtained should have been interpreted in terms of Type II error (e.g. false negatives), rather than Type I error.

Our results, shown in Tables 2 and 3, indicate that misspecification does not seem to have any consequential effect on the occurrence of Type I error, whereas statistical power is inadequate only in the case of the misspecified endogenous con-

Table 2. Type I Error Rates for Correctly and Misspecified Models

	Correctly Specified		Misspecified	
Model	Average Path	% Significant	Average Path	% Significant
Model 1A	0.002	5%	0.002	5%
Model 1B	0.002	4%	0.003	5%
Model 1C	0.001	4%	0.002	5%
Model 2A	0.001	5%	0.001	6%
Model 2B	0.001	6%	0.001	7%
Model 2C	0.001	5%	0.001	6%

Note: Values reported are average standardized paths and the proportions of significant findings for the tested paths are all replications.

Table 3. Statistical Power Results

Model 1						
	Intercorrelation = 0.10		Intercorrelation = 0.40		Intercorrelation = 0.70	
Path	Correctly Specified	Misspecified	Correctly Specified	Misspecified	Correctly Specified	Misspecified
ξ_1 to η_1	100%	100%	100%	100%	100%	100%
ξ_1 to η_3	100%	100%	100%	100%	100%	100%
η_1 to η_2	100%	100%	100%	100%	100%	100%
η_1 to η_4	100%	100%	100%	100%	100%	100%
Model 12						
	Intercorrelation = 0.10		Intercorrelation = 0.40		Intercorrelation = 0.70	
Path	Correctly Specified	Misspecified	Correctly Specified	Misspecified	Correctly Specified	Misspecified
ξ_1 to η_1	100%	98%	100%	72%	100%	27%
ξ_1 to η_3	100%	100%	100%	100%	100%	100%
η_1 to η_2	100%	100%	100%	100%	100%	100%
η_1 to η_4	100%	100%	100%	100%	100%	100%

Note: Values reported are the proportion of significant results for each path over all replications.

struct, which is consistent with the downward bias discussed before for those paths (full results for every path in both correctly and misspecified models are available in Appendixes D and E).

Limitations, Conclusions, and Directions for Future Research

Summary and Interpretation of Research Findings

The main goal of this research was to show how apparent bias in the estimation of structural relationships due to construct misspecification was mostly a function of inattention to the scale in which relationships were expressed rather than bias in the actual relationships. As a secondary objective, we sought to improve on previous work on the existence of Type I and II errors and whether their occurrence changed as a result of construct misspecification. We achieved these by outlining the relationship between the metric of latent variables and observed coefficients, and then replicating past work in order to show the different results that are obtained when considering relationships of interest in a common metric, which allows for comparison and thus estimation of the presence or absence of bias.

Results from our simulations show that, when comparing coefficients from the population model with those estimated through simulations in a common metric, the apparent bias in the estimates all but disappears. The only exception to this phenomenon is the single path from a reflectively specified latent variable to a formatively specified one, when the latter has been misspecified as reflective. For this particular case, while actual bias is less than previously reported, it remains nonetheless large and significant. The overall picture, however, is much less dire than previously thought and, for the most part, construct misspecification does not seem to result in bias in the estimates of structural parameters. Our conclusions are, of course, subject to the limitations described below.

Great care should be taken when interpreting our results. Our findings should not be taken to indicate that proper construct specification is not important—much to the contrary. To the extent that the ultimate goal of researchers is to better model the theoretical relationships underlying patterns of manifest variables, then certainly properly specifying causality relationships between components of a theory is central to such a goal. Moreover, and although not generally discussed, properly understanding the direction of causality has major implications for researchers interested in the development of interventions or manipulations of those constructs of interest. The definition of the construct itself, as well as its nomological network, are impacted when defined by one specification or the other. While we do not challenge the conclusion that misspecification is a major issue in the information systems discipline, we believe there is much to be gained from a clear and accurate perspective on the magnitude of its consequences, and why those occur.

Our research indicates, however, that as far as the magnitude of the structural relationships between latent variables in a research model are concerned, the misspecification of a formative relationship as reflective does not result in major bias. Note we are not claiming here that a researcher can merely change the direction of the relationship between latent and manifest variables and obtain the same results; rather, that the bias—that is, the average departure from the underlying population value—due to this change is limited. Consider the following results from Model 1C, shown in Figure 5.

Panel (a) in Figure 5 shows the distribution of the percentage bias in the standardized relationship between ξ_1 and η_1 in Figure 1 when ξ_1 is correctly specified as formative. Given the proper specification of the model, the distribution reflects sampling variability over 1,500 replications. Thus, researchers estimating this model with a sample of N = 500 can expect their results to be within -6 percent and +6 percent of the true value, with an average bias of 0 percent. Panel (b) in Figure 5 shows the distribution of percentage bias in the relationship when ξ_1 is improperly modeled as reflective; bias here averages about 2 percent (as also shown in Figure 2), with the majority of the replications falling in the -3 percent to +7 percent range. Finally, Panel (c) in Figure 5 shows the distribution of the difference between estimates obtained, for each replication, from the properly and misspecified models. As can be seen from the figure, these are of limited magnitude and largely confined to the 0 percent to +4 percent range. Although results from Model 1C exhibit the least amount of bias of all discussed in this article, they are nonetheless particularly important because they represent the scenario where, due to the high intercorrelation between formative indicators, these could have been misspecified as reflective and still be able to satisfactorily pass the standard used to evaluate research under that specification. Distributions for other cases are available from the first author upon request.

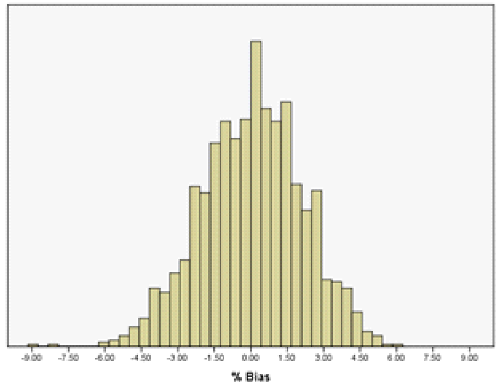
For the second objective of our research, we replicated past work on Type I and II errors and highlighted limitations in the previously used models and reporting approach that prevented a complete and accurate picture of these important issues. In particular, we contend that previous work on Type I error was conducted under the assumption that a small path between two variables was tantamount to no underlying relationship. Our decomposition of the variances involved in the models indicate that, while small, the relationship was there nonetheless. Our results show that Type I error in both correctly and misspecified models does not deviate significantly from the nominal alpha level and, therefore, construct specification does not seem to play an important role in this issue. With regard to Type II errors, the impact of construct misspecification is limited to the case described above, where downward bias in the estimate results in limited power to detect the underlying relationship, more so when the intercorrelation of the formative items is high. This is a novel finding that was not previously reported.

Limitations

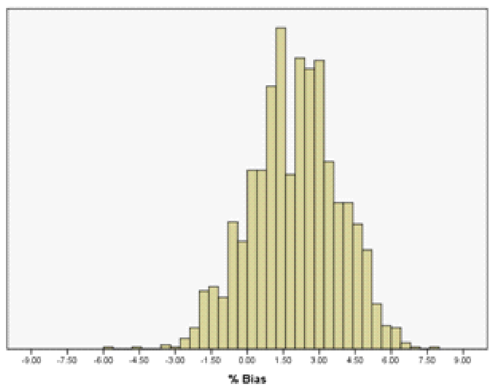
There are several limitations in this study that need to be acknowledged when interpreting these results. First, our conclusions are applicable only to the particular models studied here. While these are the most complex that have been used to make the case for the presence of bias due to misspecification (others, like those used by MacKenzie et al. [2005] included only two latent variables), our results need to be further replicated in other models with more varied structures.

Second, in the interest of replicating past work, we did not modify the magnitude of the relationships among latent variables, or any other parameter of the models. As an examination of Appendix B shows, however, the proportions of variance explained in most endogenous latent variables was somewhat high when compared to actual research. The extent to which our results would change if the relationships were more tenuous is another important limitation. This also has an impact on the generalizability of the results dealing with Type

(a) Correctly specified versus population value (standardized)



(b) Misspecified versus population value (standardized)



(c) Correctly specified versus misspecified (standardized)

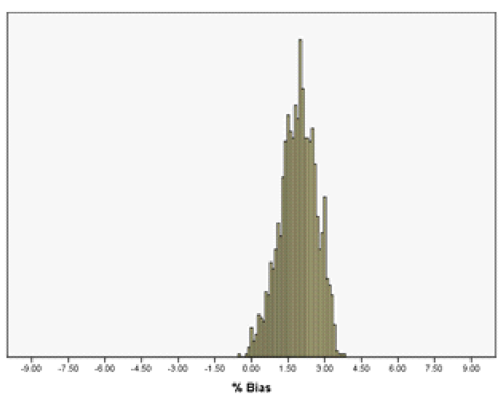


Figure 5. Distribution of Bias for Model 1C

II errors, since the magnitude of the relationship is an important determinant of statistical power such that, everything else being equal, stronger relationships are more likely to be detected than weaker ones.

Third, and also related to Type II errors, only two different sample sizes were employed in this research. A sample of 500 was used to report the results in the main body of the research, while a full replication with a sample size of 250 was included in Appendix E. The larger size was originally used by Jarvis et al. (2003), but Petter et al. (2007) noted that information systems research was generally conducted using smaller samples and chose the latter to make their results more generalizable to the discipline. Future research could also focus on the effects of varying sample size on the outcomes of interest reported here, particularly focusing on smaller samples. This, done in conjunction with effects of different magnitudes, as discussed above, would provide a more accurate picture of the occurrence of Type II error and any effects that construct misspecification has on this issue.

Fourth, our research applies only to covariance-based techniques. Other techniques based on the use of weighted aggregates of manifest variables, such as partial least squares, have also been posited to be able to accommodate formatively specified constructs. These techniques, however, do not include a disturbance term, which in the formative specification accounts for all causes of the construct not included in the model. The extent to which results from the two alternative approaches are compatible is an important issue for further examination beyond the scope of this research. While our results show that covariance-based structural equation modeling results can be misinterpreted because of scaling issues, the extent to which these findings also apply to component-based techniques, such as partial least squares, is also an open question.

Finally, our research investigated models that, aside from the particular issue under consideration here, were otherwise properly specified in all other respects. The extent to which these results hold when there are potentially other misspecifications is an area that requires further examination. For example, research by Kim et al. (2010) shows rather different standardized coefficients when alternative specifications were employed; at the same time, all models exhibited significant chi-square statistics, which may indicate the presence of other relationships not tenable in light of the collected data.

Conclusions and Future Research

The use of latent variable techniques in information systems research has been steadily increasing since they were first introduced, a trajectory likely to continue in the future given their many advantages over those dealing only with manifest variables. The recent interest in the proper specification of the relationship between latent and observed variables has challenged many of the oft-used assumptions and brought to the fore the existence of alternatives. When the ultimate objective of researchers is to build models that reflect the underlying nature of the phenomena under study, this is clearly a valuable development. It has also become a vibrant area of research, with dozens of publications in a short time span. With this work, we seek to further our understanding of the topic. Our most important contribution lies in highlighting the importance of careful consideration of the metric in which relationships are expressed, affecting whether comparisons can be performed and whether results from those are interpretable. Lack of attention to this issue underlies many of the reported results in this short but dynamic literature. Additionally, we sought to clarify the study of Type I and II errors in a manner that, we believe, is more consistent with the nature of these two issues (e.g., significance considered at the level of each individual replication). Taken together, both contributions highlight some important methodological issues that should be taken into consideration when designing and conducting further research in this area, as they have the potential to significantly impact the reporting and interpretation of the results of such studies.

In addition to those noted earlier, there are a number of valuable directions for future research. First, while much of the literature dealing with construct misspecification (Jarvis et al. 2003; MacKenzie et al. 2005; Petter et al. 2007; Podsakoff et al. 2003) is based on the possibility of researchers passing judgment solely by considering the definition of a construct and the wording of the items, it is not entirely clear that the matter is as straightforward as it first appears (Wilcox et al. 2008).

Second, while research on construct specification has focused entirely on whether the relationship should be formative or reflective, it has largely ignored another potentially important source of misspecification—that derived from omitting a formative indicator in a research model. While it is well known that reflective indicators are interchangeable, each formative one is unique and the effects of omitting one of these have not been explored. This is an important issue since researchers rarely have, particularly in the early stages of research, enough information to ensure that all possible causes of a latent variable have been accounted for and included in their models.

Third, our results show that bias appears to be an issue only in the case of endogenous latent variables with formative indicators. Although these were included in our research to enhance comparability with past literature, their conceptual interpretation is somewhat problematic. In particular, it is not clear in models such as the ones shown in Figure 1 (Model 2) how a latent variable can causally impact a formatively specified one directly without affecting any of the indicators. To the extent that the set of formative indicators and the disturbance term collectively represent all modeled and unmodeled causes of the formative latent variable, any effects on the formative latent variable should conceptually be channeled through one of those. We believe more careful consideration of the issue is warranted.

Fourth, recent developments in the SEM literature, including but not limited to the issue of formative specification of latent variables, make updated guidelines for the conduct, evaluation, and report of SEM-based research necessary. While there are a few important exemplars in this area (Boomsma 2000; Gefen et al. 2000; McDonald and Ho 2002), comprehensive treatments that include, for example, measurement invariance in multiple-group studies, new developments in the estimation and interpretation of reliability (Raykov 2004; Raykov and Penev 2006), or second order constructs are needed.

Finally, work on the development of tests of misspecification is also badly needed, particularly since past research has shown that commonly used absolute and relative fit indexes appear to be poorly equipped to detect it. Work based on confirmatory tetrad analysis (Bollen 2000) appears promising in this regard. Once these have been obtained, a more rigorous reanalysis of past research could be conducted to more precisely assess the extent of misspecification in the literature. Our understanding of many of these issues is still quite limited, and the results from these techniques may be misinterpreted even by the leading methodologists in various disciplines. Whereas, as our research shows, the consequences might not be as bleak and dire as previously thought, the proper specification of the relationships remains a worthy goal nonetheless.

References

Bagozzi, R. 2007. "On the Meaning of Formative Measurement and How it Differs from Reflective Measurement: Comment on Howell, Breivik, and Wilcox (2007)," *Psychological Methods* (12:2), pp. 229-237.

Bentler, P., and Wu, E. 1995. *EQS for Windows User's Guide*, Encino, CA: Multivariate Software, Inc.

Bollen, K. 1984. "Multiple Indicators: Internal Consistency or No Necessary Relationship," *Quality and Quantity* (18), pp. 377-385.

Bollen, K. 2000. "A Tetrad Test for Causal Indicators," *Psychological Methods* (5:1), pp. 3-22.

Bollen, K. 2007. "Interpretational Confounding Is Due to Misspecification, Not to Type of Indicator: Comment on Howell, Breivik, and Wilcox (2007)," *Psychological Methods* (12:2), pp. 219-228.

Bollen, K., and Lennox, R. 1991. "Conventional Wisdom on Measurement: A Structural Equation Perspective," *Psychological Bulletin* (110:2), pp. 305-314.

Boomsma, A. 2000. "Reporting Analyses of Covariance Structures," *Structural Equation Modeling* (7:3), pp. 461-483.

Cenfetelli, R., and Bassellier, G. 2009. "Interpretation of Formative Measurement in Information Systems Research," *MIS Quarterly* (33:4), pp. 689-708.

Cohen, P., Cohen, J., Teresi, J., Marchi, M., and Velez, C. 1990. "Problems in the Measurement of Latent Variables in Structural Equations Causal Models," *Applied Psychological Measurement* (14:2), pp. 183-196.

Curtis, R., and Jackson, E. 1962. "Multiple Indicators in Survey Research," *American Journal of Sociology* (68:2), pp. 195-204.

Diamantopoulos, A., Riefler, P., and Roth, K. 2008. "Advancing Formative Measurement Models," *Journal of Business Research* (61:12), pp. 1203-1218.

Diamantopoulos, A., and Winklhofer, H. 2001. "Index Construction with Formative Indicators: An Alternative to Scale Development," *Journal of Marketing Research* (38), pp. 269-277.

Franke, G., Preacher, K., and Rigdon, E. 2008. "Proportional Structural Effects of Formative Indicators." *Journal of Business Research* (61:12), pp. 1229-1237.

Gefen, D., Straub, D., and Boudreau, M.-C. 2000. "Structural Equation Modeling and Regression: Guidelines for Research Practice," *Communications of the Association for Information Systems* (4:7), pp. 1-77.

Gerbing, D., and Hunter, J. 1982. "The Metric of the Latent Variables in a LISREL-IV Analysis," *Educational and Psychological Measurement* (42:2), pp. 423-427.

Hardin, A., Chang, J., and Fuller, M. 2008a. "Clarifying the Use of Formative Measurement in the IS Discipline: The Case of Computer Self-Efficacy," *Journal of the Association for Information Systems* (9:9), pp. 544-546.

Hardin, A., Chang, J., and Fuller, M. 2008b. "Formative vs. Reflective Measurement: Comment on Marakas, Johnson, and Clay (2007)," *Journal of the Association for Information Systems* (9:9), pp. 519-534.

Howell, R., Breivik, E., and Wilcox, J. 2007a. "Is Formative Measurement Really Measurement? Reply to Bollen (2007) and Bagozzi (2007)," *Psychological Methods* (12:2), pp. 238-245.

Howell, R., Breivik, E., and Wilcox, J. 2007b. "Reconsidering Formative Measurement," *Psychological Methods* (12:2), pp. 205-218.

Jarvis, C., MacKenzie, S., and Podsakoff, P. 2003. "A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research," *Journal of Consumer Research* (30), pp. 199-218.

Kim, G., Shin, B., and Grover, V. 2010. "Investigating Two Contradictory Views of Formative Measurement in Information Systems Research," *MIS Quarterly* (34:2), pp. 345-365.

Kline, R. 2005. *Principles and Practice of Structural Equation Modeling*, New York: The Guilford Press.

MacKenzie, S., Podsakoff, P., and Jarvis, C. 2005. "The Problem of Measurement Model Misspecification in Behavioral and Organizational Research and Some Recommended Solutions," *Journal of Applied Psychology* (90:4), pp. 710-730.

Marakas, G., Johnson, R., and Clay, P. 2007. "The Evolving Nature of the Computer Self-Efficacy Construct: An Empirical Investigation of Measurement Construction, Validity, Reliability and Stability Over Time," *Journal of the Association for Information Systems* (8:1), pp. 16-46.

Marakas, G., Johnson, R., and Clay, P. 2008. "Formative vs. Reflective Measurement: A Reply to Hardin, Chang, and Fuller," *Journal of the Association for Information Systems* (9:9), pp. 535-543.

McDonald, R., and Ho, M.-H. 2002. "Principles and Practice in Reporting Structural Equation Analyses," *Psychological Methods* (7:1), pp. 64-82.

Millsap, R. 2001. "When Trivial Constraints Are Not Trivial: The Choice of Uniqueness Constraints in Confirmatory Factor Analysis," *Structural Equation Modeling* (8:1), pp. 1-17.

Petter, S., Straub, D., and Rai, A. 2007. "Specifying Formative Constructs in Information Systems Research," *MIS Quarterly* (31:4), pp. 623-656.

Podsakoff, P., MacKenzie, S., Podsakoff, N., and Lee, J. 2003. "The Mismeasure of Man(agement) and its Implications for Leadership Research," *The Leadership Quarterly* (14), pp. 615-656.

Raykov, T. 2004. "Estimation of Maximal Reliability: A Note on a Covariance Structure Modelling Approach," *British Journal of Mathematical and Statistical Psychology* (57), pp. 21-27.

Raykov, T., and Penev, S. 2006. "A Direct Method for Obtaining Approximate Standard Error and Confidence Interval of Maximal Reliability for Composites with Congeneric Measures," *Multivariate Behavioral Research* (4:1), pp. 15-28.

Steiger, J. 2002. "When Constraints Interact: A Caution About Reference Variables, Identification Constraints, and Scale Dependencies in Structural Equation Modeling," *Psychological Methods* (7:2), pp. 210-227.

Wilcox, J., Howell, R., and Breivik, E. 2008. "Questions About Formative Measurement," *Journal of Business Research* (61:12), pp. 1219-1228.

About the Authors

Miguel I. Aguirre-Urreta is an assistant professor of Information Systems at the School of Accountancy and Management Information Systems, College of Commerce, at DePaul University. He received his Ph.D. in Information Systems from the University of Kansas and his MBA from the Kelley School of Business at Indiana University. His teaching experience includes introduction to information systems, accounting information systems, management of information technology, and accounting information systems and auditing. His research interests include computer self-efficacy, technology acceptance, quantitative research methods, and the use of simulation in both research and practice. His work has previously appeared in *The DATA BASE for Advances in Information Systems*, *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, and *Journal of Organizational and End User Computing,* as well as numerous national and international conferences.

George M. Marakas is a professor of Information Systems at the School of Business at the University of Kansas. He received his Ph.D. in Information Systems from Florida International University in Miami and his MBA from Colorado State University. His teaching expertise includes systems analysis and design, technology-assisted decision making, electronic commerce, management of IS resources, behavioral IS research methods, and data visualization and decision support. In addition, George is an active researcher in the area of systems analysis methods, data mining and visualization, creativity enhancement, conceptual data modeling, technology acceptance, and computer self-efficacy. He has received numerous national teaching awards and his research has appeared in the top journals in his field. He is also the author of five best-selling textbooks in information systems.

MIS Quarterly

RESEARCH NOTE

REVISITING BIAS DUE TO CONSTRUCT MISSPECIFICATION: DIFFERENT RESULTS FROM CONSIDERING COEFFICIENTS IN STANDARDIZED FORM

Miguel I. Aguirre-Urreta
School of Accountancy and MIS, College of Commerce, DePaul University,
1 East Jackson Boulevard, Chicago, IL 60604 U.S.A. {maguirr6@depaul.edu}

George M. Marakas
School of Business, University of Kansas, Summerfield Hall,
1300 Sunnyside Avenue, Lawrence, KS 66045 U.S.A. {gmarakas@ku.edu}

Appendix A

Effect of Scale Metric on Standardized Results (Example)

We illustrate the importance of this issue with a simple example. Following the conventions set by Wright (1934), all models depicted in this research show causal influence by unidirectional arrows from cause to effect, and relationships not analyzed in causal terms (correlations, covariances) by curved two-headed arrows. From the population model shown in Figure A1, we generated one dataset with 100,000 observations, and then fit this data to models with identification constraints differing on both approach (reference loading or variance of the exogenous latent variable) and nonzero values (1, 0.50, and 2) yielding a total of 18 possible combinations. The size of the generated dataset was chosen not for its representativeness in applied research, but rather to ensure the estimated parameters would be virtually free from sampling error and thus would remove this source of influence from the point we are trying to make. The population value for the standardized path was 0.630 and the proportion of variance explained in the endogenous variable 39.5 percent. Results from this exercise are shown in Table A1.

Table A1. Results from Alternative Identification Constraints

Identification of Exogenous Variable	Identification of Endogenous Variable								
	Loading = 1			Loading = 0.5			Loading = 2		
	Unstd.	Std.	χ^2	Unstd.	Std.	χ^2	Unstd.	Std.	χ^2
Loading = 1	0.879	0.630	3.605	1.759	0.630	3.605	0.440	0.630	3.605
Loading = 0.5	0.440	0.630	3.605	0.879	0.630	3.605	0.220	0.630	3.605
Loading = 2	1.759	0.630	3.605	3.518	0.630	3.605	0.879	0.630	3.605
Variance = 1	1.491	0.630	3.605	2.983	0.630	3.605	0.746	0.630	3.605
Variance = 0.5	2.109	0.630	3.605	4.218	0.630	3.605	1.055	0.630	3.605
Variance = 2	1.055	0.630	3.605	2.109	0.630	3.605	0.527	0.630	3.605

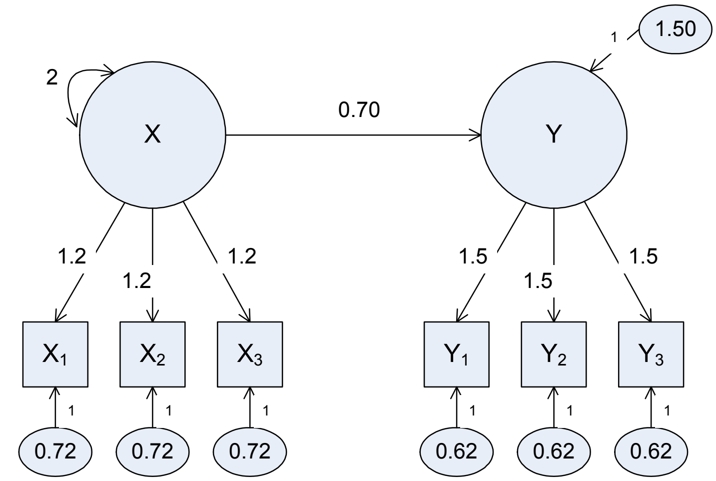


Figure A1. Population Model for Identification Example

Before attempting to interpret these results, it should be noted that all models are correctly specified, as follows: the direction of causality between the latent variables matches the one at the population level, all observed variables load on their respective factors, and the direction of these relationships is also correct. Therefore, all parameter estimates obtained should be unbiased with respect to the corresponding population values, aside from the effects of sampling error. Table A1, however, shows unstandardized estimates that are markedly different from each other, and from known population values. As shown by the χ^2 value, however, model fit is identical across models with different identification constraints, as are the respective standardized path coefficients, from which it follows that the proportion of variance explained in the endogenous latent variable is also the same across all conditions.

The interpretation of these results is quite straightforward, and entirely consistent with our discussion of identification and latent variable metric issues. Whereas unstandardized regression coefficients are a function of the particular approach and particular nonzero values used to set the scale for the latent variables, the nature of the underlying relationship between the two latent variables remains unchanged as a result of these choices. The appearance of bias in the estimate is the result of having used values to set the scales of the latent variables that differ from those at the population level which, incidentally, would be unknown to a researcher attempting to estimate these models. Moreover, even if we had restricted our choice of values to the commonly used unity, we would still have seen apparent bias in the estimate, as evidenced by the result shown in the top-left cell in Table A1 (where the unstandardized regression estimate of 0.879 clearly differs from the population value of 0.700). Other researchers (e.g., Marsh et al. 2004) have recognized the importance of this issue and have rescaled items accordingly to avoid the confounding that would arise from interpreting results expressed in different metrics.

Unstandardized estimates, both regression coefficients as well as covariances between latent variables, are expressed in a metric that is dependent on the particular estimates of the variances for the two latent variables involved in the relationship, and cannot be easily interpreted as to their magnitude absent knowledge about those. In this particular example, the smallest unstandardized estimate was 0.220 (from the combination of a loading set at a value of 2 in the endogenous variable, and at a value of 0.50 in the exogenous one), and the largest 4.218 (from the combination of a loading in the endogenous variable set at 0.50 and the variance of the exogenous latent variable also fixed at 0.50), yet both represent the same proportion of variance explained in the endogenous factor by the exogenous one when the variances of each variable are taken into account. In the first case, the estimated variance of the exogenous variable was 11.503 and that of the endogenous one 1.340, for an R^2 of 39.77 percent [$(11.503 \times 0.220^2) / 1.340$; see Appendix B for a discussion on how to decompose the variances in latent variable models], not different from the population R^2 of 39.5 percent. In the second case, the variance of the exogenous variable was estimated at 0.50 (by virtue of being fixed at that value) and the variance for the endogenous one at 22.380, for an R^2 of 39.75 percent [$(0.50 \times 4.218^2) / 22.380$].

On the other hand, while unstandardized coefficients are a function of the estimated variances of the involved variables, standardized coefficients are obtained by rescaling the research model so that all involved variables have a variance of one. All commonly used statistical packages (e.g., LISREL, MPlus, EQS, etc.) provide these estimates. As shown in Table A1, the standardized regression coefficients were always the same across all conditions, and matched their population level counterparts, as did the proportions of variance explained in the endogenous variable across all conditions.

Appendix B

Decomposition of Variances in the Research Models

The formulas employed in this appendix are

1. The variance of the sum of uncorrelated variables equals the sum of their variances, such that VAR (X + Y) = VAR (X) + VAR (Y).

2. The variance of the sum of correlated variables equals the sum of their variances plus two times their covariance, such that VAR (X + Y) = VAR (X) + 2 COV (X , Y) + VAR (Y).

3. The variance of a random variable multiplied by a constant equals the variance of the random variable times the square of the constant, such that VAR (cX) = c^2 VAR (X).

The main contention of this research is that the scale in which results are expressed has not been carefully considered, and that has led to equivocal conclusions about the different effects discussed here. One example of these issues is the statement by Jarvis et al. (2003) that the models shown in Figure 1 were built so that the item error variances would average 32 percent per factor, consistent with the average amount of random and systematic error found in marketing studies, and thus making these models more representative of that literature. This figure, 32 percent, arises from averaging the residual variances of each of the four items loading on the reflectively specified factors in Figure 1, namely 0.40, 0.24, 0.34, and 0.30 (see, for example, those for η_1 in Models 1A, 1B, and 1C). An average error of 0.32 would represent an average 32 percent of the variance in these indicators only if the average indicator variance amounted to one. Our decomposition of the full and explained variances for all correctly specified models depicted in Figure 1 shows this is not the case. Rather, the average item error variance for η_1 and η_3 in Model 1A above comes to 24 percent and 27 percent respectively and decreases as the formative item intercorrelation increases, reaching 16 percent and 22 percent, respectively, when those are 0.70 (i.e., Model 1C).

This decrease occurs as follows. As the item intercorrelation for the formatively specified construct increases, so does the variance of that construct. This in turn leads to an increase in the variance of those constructs on the receiving end of a path emitting from the formatively specified construct, which in turn results in an increase in the variance for their indicators. Since the residual variance is fixed and does not change from model to model, the proportion this residual item variance represents of the overall total decreases accordingly. Similar results arise for all other correctly specified models, being more markedly different for those in Models 2A, 2B, and 2C. The decomposition shown below also reveals that the proportion of variance explained in the constructs themselves is rather high, ranging from 42 percent to 70 percent in the reflectively specified constructs in Models 1A, 1B, and 1C, and from 17 percent to 71 percent in the reflectively specified constructs in Models 2A, 2B, and 2C above.

We next show how to calculate both the variance and the proportion of variance explained for each variable in all four models. When the formatively specified construct only emits paths to other latent variables (e.g., Models 1A, 1B, and 1C in Figure 1), the equations relating observed with latent variables, and latent variables among themselves are (intercepts omitted for clarity, as they do not affect the calculations below)

$$\xi_1 = \gamma_1 X_1 + \gamma_2 X_2 + \gamma_3 X_3 + \gamma_4 X_4 + \zeta_0$$
$$\eta_1 = \beta_1 \xi_1 + \zeta_1$$
$$\eta_3 = \beta_3 \xi_1 + \zeta_3$$
$$\eta_2 = \beta_2 \eta_1 + \zeta_2$$
$$\eta_4 = \beta_4 \eta_1 + \zeta_4$$
$$y_i = \lambda_i \eta_j + \delta_i$$

Therefore, and using the formulas presented above, the following decomposition of the variances for each observed and latent variable can be made, using Model 1A as an example. First, the variance of the formative variable equals (all gamma coefficients were set to one in the population model used in the simulations) $VAR(\xi_1) = VAR(1X_1 + 1X_2 + 1X_3 + 1X_4 + \zeta_0)$. Since the disturbance term ζ_0 is uncorrelated with the exogenous variables X_{1-4}, then (by formula 1 above) $VAR(1X_1 + 1X_2 + 1X_3 + 1X_4 + \zeta_0) = VAR(1X_1 + 1X_2 + 1X_3 + 1X_4) + VAR(\zeta_0)$. The X variables, however, are correlated among themselves, thus (by formula 2 above and omitting the ones)

$$VAR(1X_1 + 1X_2 + 1X_3 + 1X_4) + VAR(\zeta_0) = VAR(X_1) + VAR(X_2) + VAR(X_3) + VAR(X_4) + 2\,COV(X_1, X_2) + 2\,COV(X_1, X_3) + 2\,COV(X_1, X_4) + 2\,COV(X_2, X_3) + 2\,COV(X_2, X_4) + 2\,COV(X_3, X_4) + VAR(\zeta_0)$$

Replacing with the population values from Model 1A (e.g., the variances of the X variables equal 1, their correlation equals 0.10, and the variance of the disturbance term, that is, the residual variance, equals 0.50)

$$VAR(\xi_1) = 1 + 1 + 1 + 1 + 6 \times 2 \times 0.10 + 0.50 = 5.70$$

Therefore, the proportion of the variance of ξ_1 explained by X_1, X_2, X_3, and X_4 is 91.23 percent [(5.70 – 0.50) / 5.70]. Next, since η_1 and η_3 are related to ξ_1 by the same regression parameter, their variances will be the same; for η_1 $VAR(\eta_1) = VAR(\beta_1\xi_1 + \zeta_1)$; by formulas 1 and 3 above, the decomposition becomes $VAR(\beta_1\xi_1 + \zeta_1) = \beta_1^2 VAR(\xi_1) + VAR(\zeta_1)$, and replacing with the population parameters (0.30 for the beta coefficient, 5.70 for the variance of the formative variable as calculated above, and 0.50 for the variance of the disturbance term) obtains $VAR(\eta_1) = 0.30^2 \times 5.70 + 0.50 = 1.013$. Therefore, the proportion of the variance of η_1 explained by ξ_1 equals 50.64 percent [(1.013 – 0.50) / 1.013]. The same values apply to η_3.

The variances of all other variables in the model (latent and manifest) can be similarly obtained. Table B1 shows these values, together with the proportions of variance explained in each, for Models 1A, 1B, and 1C in Figure 1 (which differ only in terms of the correlation among the cause indicators of the formatively specified construct).

Table B1. Variances of Variables

	Model 1A (Correlations = 0.10)		Model 1B (Correlations = 0.40)		Model 1C (Correlations = 0.70)	
Variable	Variance	R^2	Variance	R^2	Variance	R^2
ξ_1	5.700	91.23%	9.300	94.62%	12.900	96.12%
η_1, η_3	1.013	50.64%	1.337	62.60%	1.661	69.90%
η_2, η_4	0.865	42.20%	0.981	49.03%	1.098	54.46%
Y_1, Y_9	1.413	71.69%	1.737	76.97%	2.061	80.60%
Y_2, Y_{10}	1.253	80.85%	1.577	84.78%	1.901	87.38%
Y_3, Y_{11}	1.353	74.87%	1.677	79.73%	2.001	83.00%
Y_4, Y_{12}	1.313	77.15%	1.637	81.67%	1.961	84.70%
Y_5, Y_{13}	1.265	68.38%	1.381	71.04%	1.498	73.30%
Y_6, Y_{14}	1.105	78.28%	1.221	80.34%	1.338	82.06%
Y_7, Y_{15}	1.205	71.78%	1.321	74.26%	1.438	76.36%
Y_8, Y_{16}	1.165	74.25%	1.281	76.58%	1.398	78.54%

The proportion of variance explained and, by subtraction, the residual item variance for each indicator loading on a reflectively specified factor can be obtained from Table B1 as follows: For those indicators loading on η_1 and η_3 ($Y_{1–4}$ and $Y_{9–12}$), the average variance explained is 76.14 percent [(71.69 + 80.85 + 74.87 + 77.15) / 4] in Model 1A, 80.79 percent in Model 1B, and 83.92 percent in Model 1C; therefore, the average item error variance for these indicators is 23.86 percent, 19.21 percent, and 16.08 percent for Models 1A, 1B, and 1C respectively. For indicators loading on η_2 and η_4 ($Y_{5–8}$ and $Y_{13–16}$), the corresponding values are 73.17 percent, 75.56 percent, and 77.57 percent for the proportion of variance explained, and 26.83 percent, 24.45 percent, and 22.44 percent for the average item error variance, for Models 1A, 1B, and 1C, respectively.

Using the same formulas and the appropriate equations relating observed to latent variables, as well as latent variables among themselves, Table B2 shows variances and the proportions explained in each, for Models 2A, 2B, and 2C in Figure 1 (differing in terms of the correlation among the cause indicators of the formatively specified construct).

Table B2. Variances and Proportions Explained

Variable	Model 2A (Correlations = 0.10)		Model 2B (Correlations = 0.40)		Model 2C (Correlations = 0.70)	
	Variance	R^2	Variance	R^2	Variance	R^2
ξ_1	0.104	—	0.104	—	0.104	—
η_1	6.636	92.46%	10.236	95.11%	13.836	96.39%
η_3	0.604	17.22%	0.604	17.22%	0.604	17.22%
η_2, η_4	1.097	54.43%	1.421	64.82%	1.745	71.35%
X_1	0.504	20.63%	0.504	20.63%	0.504	20.63%
X_2	0.344	30.23%	0.344	30.23%	0.344	30.23%
X_3	0.444	23.42%	0.444	23.42%	0.444	23.42%
X_4	0.404	25.74%	0.404	25.74%	0.404	25.74%
Y_9	1.004	60.16%	1.004	60.16%	1.004	60.16%
Y_{10}	0.844	71.56%	0.844	71.56%	0.844	71.56%
Y_{11}	0.944	63.98%	0.944	63.98%	0.944	63.98%
Y_{12}	0.904	66.81%	0.904	66.81%	0.904	66.81%
Y_5, Y_{13}	1.497	73.28%	1.821	78.04%	2.145	81.35%
Y_6, Y_{14}	1.337	82.05%	1.661	85.55%	1.985	87.91%
Y_5, Y_{15}	1.437	76.34%	1.761	80.70%	2.085	83.69%
Y_8, Y_{16}	1.397	78.53%	1.721	82.57%	2.045	85.33%

For the manifest variables loading on ξ_1 ($X_{1–4}$) the average proportion of explained variance in them by this factor was 25.01 percent in all three cases, as the variance of this factor did not change from model to model. This result in the average item error for these indicators is74.99 percent, quite different from the 32 percent stated by Jarvis et al. (2003). For manifest variables loading on η_2 and η_4 ($Y_{5–8}$ and $Y_{13–16}$), the average variance explained is 77.55 percent, 81.72 percent, and 84.57 percent for Models 2A, 2B, and 2C, yielding an average item error variance of 22.45 percent, 18.29 percent, and 15.43 percent, respectively.

Finally, the items loading on η_3 also display the same explained and residual variance from model to model, since changes in the intercorrelation of the items for the formatively specified constructs had no effect on them. The average variance explained in these indicators amounts to 65.63 percent and the average item residual to 34.37 percent.

Appendix C

Calculation of Standardized Path Coefficients

The calculation of standardized values for all population parameters in the simulated models shown in Figure 1 is quite straightforward and can be performed with any common SEM software package by fixing all parameter estimates at their population values and applying the resulting model to any covariance matrix with the same number of observed variables (since most packages do require some input, even if it is not used in any meaningful way since models contain only fixed values; in our case we used an identity matrix). The following code uses MPlus 3.0 to estimate Model 1A:

```
TITLE: Model 1A Population Values
DATA:
FILE IS C:\I-MATRIX20V.TXT;
TYPE IS COVARIANCE;
NOBSERVATIONS ARE 100000;
VARIABLE:
NAMES ARE x1-x4 y1-y16;
ANALYSIS:
TYPE = GENERAL;
ESTIMATOR = ML;
MODEL:
E1@0.50
E1 BY y1-y4@1;
y1@0.40 y2@0.24 y3@0.34 y4@0.30;
E3@0.50
E3 BY y9-y12@1;
y9@0.40 y10@0.24 y11@0.34 y12@0.30;
K1 BY E1@0.30 E3@0.30;
K1 ON x1-x4@1;
K1@0.50;
x1 WITH x2@0.10 x3@0.10 x4@0.10;
x2 WITH x3@0.10 x4@0.10;
x3 WITH x4@0.10;
x1-x4@1;
E4@0.50
E4 BY y13-y16@1;
y13@0.40 y14@0.24 y15@0.34 y16@0.30;
E2@0.50
E2 BY y5-y8@1;
y5@0.40 y6@0.24 y7@0.34 y8@0.30;
E2 ON E1@0.60;
E4 on E1@0.60;
E4 WITH K1@0;
E2 WITH K1@0;
E2 WITH E4@0;
OUTPUT: RESIDUAL STAND;
```

Running this analysis provides the researcher, among other things, with the population values in standardized metric for each path or parameter of interest. For Model 1A, the relevant standardized values are 0.712 for the K → E1 and K → E3 paths, and 0.649 for the E1 → E2 and E1 → E4 paths. Generally, Table C1 shows the standardized population values for all models, and their variations, investigated in this research. It should be noted that these values for Models 3 and 4 are identical to their counterparts in Models 1 and 2 since the former test for the presence of a nonexistent relationship, which does not alter the population values for these paths.

Table C1. Standardized Population Values

Model	K → E1	K → E3	E1 → E2	E1 → E4
Model 1A (formative indicators correlated at 0.10)	0.712	0.712	0.649	0.649
Model 1B (formative indicators correlated at 0.40)	0.791	0.791	0.700	0.700
Model 1C (formative indicators correlated at 0.70)	0.836	0.839	0.738	0.738
Model 2A (formative indicators correlated at 0.10)	0.376	0.415	0.738	0.738
Model 2B (formative indicators correlated at 0.40)	0.302	0.415	0.805	0.805
Model 2C (formative indicators correlated at 0.70)	0.260	0.415	0.845	0.845

Note: These models and paths correspond to those shown in Figure 1.

It may seem puzzling that, while the structural coefficients in the three variations of both Models 1 and 2 (see Figure 1) do not change, their standardized counterparts do, as shown in Table C1. This is, however, a direct result of the way formatively specified latent variables perform in these models. We work out one example in more detail to show why this is the case.

Consider the relationship between ξ_1 and η_1 in Model 1. The unstandardized coefficient in this relationship is always 0.300 in all variations shown in Figure 1, yet its standardized version, as shown in Table C1, varies from Model 1A to 1B to 1C. We emphasize, however, that the nature of the relationship does not change, as follows: From Appendix B, the variances of ξ_1 and η_1 in Model 1A are 5.700 and 1.013, respectively. Since the structural coefficient linking both latent variables is 0.300 (here, all numbers are expressed in unstandardized form), the variance explained in η_1 by ξ_1 is $5.700 \times 0.300^2 = 0.513$ or $0.513 / 1.013 = 0.5064$ or 50.64 percent. Following the same logic and drawing from the calculations in Appendix B, the proportion of variance explained is 62.60 percent in Model 1B and 69.90 percent in Model 1C.

Recall that when coefficients are standardized they are expressed to show a relationship when the variances of the variables involved are one; for example, standardizing a covariance results in a correlation between two variables. In Table C1, the standardized coefficients for this relationship are 0.712, 0.791, and 0.836 for Models 1A, 1B, and 1C, respectively. Since standardized coefficients involve variances equal to one, the proportions of variance explained for each model are $0.712^2 = 0.5069$ or 50.69 percent, $0.791^2 = 0.6257$ or 62.57 percent, and $0.836^2 = 0.6989$ or 69.89 percent. Minor differences with the percentages shown above are due to rounding of the standardized coefficients to three decimal places.

Therefore, although the structural coefficients, in unstandardized form, remain the same across all three variations of each model shown in Figure 1, the variances of the variables involved do not, and thus the same unstandardized coefficient represents a different relationship in each instance. When expressed in standardized form, these coefficients are different in each case so that the relationship can stay the same, as shown above.

Appendix D

Results from Simulations with N = 500

Table D1. Average Standardized and Unstandardized Paths, and Deviations from Population Values (Models 1 and 2, over 1,500 replications)

Path	Correctly Specified Models				Misspecified Models			
	Unstd.	% Dev.	Std.	% Dev.	Unstd.	% Dev.	Std.	% Dev.
Model 1A (formative indicators correlated at 0.10)								
K → E1	0.299	-0.3%	0.712	0.0%	1.762	487.3%	0.763	7.2%
K → E3	0.299	-0.3%	0.712	0.0%	1.754	484.7%	0.759	6.6%
E1 → E2	0.600	0.0%	0.649	0.0%	0.600	0.0%	0.649	0.0%
E1 → E4	0.600	0.0%	0.649	0.0%	0.600	0.0%	0.649	0.0%
Model 1B (formative indicators correlated at 0.40)								
K → E1	0.299	-0.3%	0.792	0.1%	1.456	385.3%	0.827	4.6%
K → E3	0.299	-0.3%	0.792	0.1%	1.454	384.7%	0.827	4.3%
E1 → E2	0.600	0.0%	0.700	0.0%	0.600	0.0%	0.700	0.0%
E1 → E4	0.600	0.0%	0.700	0.0%	0.600	0.0%	0.700	0.0%
Model 1C (formative indicators correlated at 0.70)								
K → E1	0.298	-0.7%	0.836	0.0%	1.310	336.7%	0.852	1.9%
K → E3	0.298	-0.7%	0.836	0.0%	1.310	336.7%	0.852	1.9%
E1 → E2	0.600	0.0%	0.738	0.3%	0.600	0.0%	0.738	0.3%
E1 → E4	0.600	0.0%	0.738	0.3%	0.600	0.0%	0.738	0.3%
Model 2A (formative indicators correlated at 0.10)								
K → E1	3.071	2.4%	0.377	0.3%	0.350	-88.3%	0.265	-29.5%
K → E3	1.012	1.2%	0.416	0.2%	1.013	1.3%	0.416	0.2%
E1 → E2	0.299	-0.3%	0.737	-0.1%	1.950	550.0%	0.776	5.1%
E1 → E4	0.298	-0.7%	0.737	-0.1%	1.949	549.7%	0.777	5.3%
Model 2B (formative indicators correlated at 0.40)								
K → E1	3.089	3.0%	0.303	0.3%	0.331	-89.0%	0.164	-45.7%
K → E3	1.012	1.2%	0.416	0.2%	1.013	1.3%	0.416	0.2%
E1 → E2	0.298	-0.7%	0.804	-0.1%	1.531	410.3%	0.828	2.9%
E1 → E4	0.298	-0.7%	0.805	0.0%	1.528	409.3%	0.828	2.9%
Model 2C (formative indicators correlated at 0.70)								
K → E1	3.135	4.5%	0.261	0.4%	0.209	-93.0%	0.080	-69.2%
K → E3	1.012	1.2%	0.416	0.2%	1.013	1.3%	0.416	0.2%
E1 → E2	0.297	-1.0%	0.844	-0.1%	1.335	345.0%	0.843	-0.2%
E1 → E4	0.297	-1.0%	0.844	-0.1%	1.331	343.7%	0.843	-0.2%

Note: Unstd. = average of unstandardized paths over replications.
Std. = average of standardized paths over replications.
% Dev. = average deviation from population value, calculated as (average of paths – population value) / population value.

Table D2. Average Standardized and Unstandardized Paths, and Deviations from Population Values (Models 3 and 4, over 1,500 replications)								
	Correctly Specified Models				Misspecified Models			
Path	Unstd.	% Dev.	Std.	% Dev.	Unstd.	% Dev.	Std.	% Dev.
Model 3A (formative indicators correlated at 0.10)								
K → E1	0.299	-0.3%	0.712	0.0%	1.764	488.0%	0.764	7.3%
K → E3	0.299	-0.3%	0.712	0.0%	1.752	484.0%	0.759	6.6%
E1 → E2	0.600	0.0%	0.649	0.0%	0.600	0.0%	0.649	0.0%
E1 → E4	0.599	-0.2%	0.648	-0.2%	0.599	-0.2%	0.647	-0.3%
K → E4	0.001		0.002		0.005		0.002	
Model 3B (formative indicators correlated at 0.40)								
K → E1	0.299	-0.3%	0.792	0.1%	1.456	385.3%	0.826	4.4%
K → E3	0.298	-0.7%	0.792	0.1%	1.454	384.7%	0.825	4.3%
E1 → E2	0.600	0.0%	0.700	0.0%	0.600	0.0%	0.700	0.0%
E1 → E4	0.599	-0.2%	0.698	-0.3%	0.598	-0.3%	0.698	-0.3%
K → E4	0.001		0.002		0.004		0.003	
Model 3C (formative indicators correlated at 0.70)								
K → E1	0.298	-0.7%	0.836	0.0%	1.310	336.7%	0.852	1.9%
K → E3	0.298	-0.7%	0.836	0.0%	1.310	336.7%	0.852	1.9%
E1 → E2	0.600	0.0%	0.738	0.3%	0.600	0.0%	0.738	0.3%
E1 → E4	0.600	0.0%	0.737	0.1%	0.599	-0.2%	0.737	0.1%
K → E4	0.000		0.001		0.002		0.002	
Model 4A (formative indicators correlated at 0.10)								
K → E1	3.070	2.3%	0.376	0.0%	0.349	-88.4%	0.265	-29.5%
K → E3	1.010	1.0%	0.415	0.0%	1.014	1.4%	0.416	0.2%
E1 → E2	0.299	-0.3%	0.737	-0.1%	1.950	550.0%	0.776	5.1%
E1 → E4	0.298	-0.7%	0.737	-0.1%	1.948	549.3%	0.777	5.3%
E1 → E3	0.000		0.001		0.002		0.001	
Model 4B (formative indicators correlated at 0.40)								
K → E1	3.088	2.9%	0.303	0.3%	0.331	-89.0%	0.164	-45.7%
K → E3	1.011	1.1%	0.415	0.0%	1.014	1.4%	0.416	0.2%
E1 → E2	0.298	-0.7%	0.805	0.0%	1.531	410.3%	0.828	2.9%
E1 → E4	0.298	-0.7%	0.805	0.0%	1.527	409.0%	0.828	2.9%
E1 → E3	0.000		0.001		0.002		0.001	
Model 4C (formative indicators correlated at 0.70)								
K → E1	3.135	4.5%	0.261	0.4%	0.208	-93.1%	0.080	-69.2%
K → E3	1.011	1.1%	0.415	0.0%	1.013	1.3%	0.416	0.2%
E1 → E2	0.297	-1.0%	0.844	-0.1%	1.335	345.0%	0.843	-0.2%
E1 → E4	0.297	-1.0%	0.844	-0.1%	1.331	343.7%	0.843	-0.2%
E1 → E3	0.000		0.001		0.001		0.001	

Note: Unstd. = average of unstandardized paths over replications.
Std. = average of standardized paths over replications.
% Dev. = average deviation from population value, calculated as (average of paths – population value) / population value. Not calculated for the K → E4 path, which is zero at the population value.

Table D3. Statistical Power Results ((Models 1 and 2, over 1,500 replications)

Model	K → E1	K → E3	E1 → E2	E1 → E4
Model 1A (indicators correlated at 0.10)				
Correctly specified	100.0%	100.0%	100.0%	100.0%
Misspecified	100.0%	100.0%	100.0%	100.0%
Model 1B (indicators correlated at 0.40)				
Correctly specified	100.0%	100.0%	100.0%	100.0%
Misspecified	100.0%	100.0%	100.0%	100.0%
Model 1C (indicators correlated at 0.70)				
Correctly specified	100.0%	100.0%	100.0%	100.0%
Misspecified	100.0%	100.0%	100.0%	100.0%
Model 2A (indicators correlated at 0.10)				
Correctly specified	100.0%	100.0%	100.0%	100.0%
Misspecified	97.6%	100.0%	100.0%	100.0%
Model 2B (indicators correlated at 0.40)				
Correctly specified	100.0%	100.0%	100.0%	100.0%
Misspecified	72.1%	100.0%	100.0%	100.0%
Model 2C (indicators correlated at 0.70)				
Correctly specified	100.0%	100.0%	100.0%	100.0%
Misspecified	27.1%	100.0%	100.0%	100.0%

Table D4. Statistical Power and Type I Error Results (Model 3, over 1,500 replications)

Model	K → E1	K → E3	E1 → E2	E1 → E4	K → E4
Model 3A (indicators correlated at 0.10)					
Correctly specified	100.0%	100.0%	100.0%	100.0%	4.5%
Misspecified	100.0%	100.0%	100.0%	100.0%	4.6%
Model 3B (indicators correlated at 0.40)					
Correctly specified	100.0%	100.0%	100.0%	100.0%	4.0%
Misspecified	100.0%	100.0%	100.0%	100.0%	4.8%
Model 3C (indicators correlated at 0.70)					
Correctly specified	100.0%	100.0%	100.0%	100.0%	3.7%
Misspecified	100.0%	100.0%	100.0%	100.0%	4.8%

Note: For the K → E4 path, reported values represent Type I error occurrence.

Table D5. Statistical Power and Type I Error Results (Model 4, over 1,500 replications)

Model	K → E1	K → E3	E1 → E2	E1 → E4	E1 → E3
Model 4A (indicators correlated at 0.10)					
Correctly specified	100.0%	100.0%	100.0%	100.0%	5.7%
Misspecified	96.9%	100.0%	100.0%	100.0%	6.2%
Model 4B (indicators correlated at 0.40)					
Correctly specified	100.0%	100.0%	100.0%	100.0%	5.5%
Misspecified	70.9%	100.0%	100.0%	100.0%	6.8%
Model 4C (indicators correlated at 0.70)					
Correctly specified	100.0%	100.0%	100.0%	100.0%	4.8%
Misspecified	25.7%	100.0%	100.0%	100.0%	6.1%

Note: For the E1 → E3 path, reported values represent Type I error occurrence.

Appendix E

Results from Simulations with N = 250

Table E1. Average Standardized and Unstandardized Paths, and Deviations from Population Values (Models 1 and 2, over 1,500 replications)

	Correctly Specified Models				Misspecified Models			
Path	Unstd.	% Dev.	Std.	% Dev.	Unstd.	% Dev.	Std.	% Dev.
Model 1A (formative indicators correlated at 0.10)								
K → E1	0.300	0.0%	0.711	-0.1%	1.769	489.7%	0.759	6.6%
K → E3	0.300	0.0%	0.713	0.1%	1.774	491.3%	0.761	6.9%
E1 → E2	0.601	0.2%	0.647	-0.3%	0.601	0.2%	0.647	-0.3%
E1 → E4	0.603	0.5%	0.649	0.0%	0.603	0.5%	0.649	0.0%
Model 1B (formative indicators correlated at 0.40)								
K → E1	0.301	0.3%	0.791	0.0%	1.456	385.3%	0.825	4.3%
K → E3	0.301	0.3%	0.792	0.1%	1.456	385.3%	0.825	4.3%
E1 → E2	0.600	0.0%	0.698	-0.3%	0.600	0.0%	0.698	-0.3%
E1 → E4	0.602	0.3%	0.700	0.0%	0.602	0.3%	0.700	0.0%
Model 1C (formative indicators correlated at 0.70)								
K → E1	0.302	0.7%	0.836	0.0%	1.311	337.0%	0.851	1.8%
K → E3	0.302	0.7%	0.836	0.0%	1.310	336.7%	0.852	1.9%
E1 → E2	0.600	0.0%	0.736	0.0%	0.602	0.3%	0.736	0.0%
E1 → E4	0.602	0.3%	0.738	0.3%	0.600	0.0%	0.738	0.3%
Model 2A (formative indicators correlated at 0.10)								
K → E1	3.187	6.2%	0.377	0.3%	0.361	-88.0%	0.267	-29.0%
K → E3	1.030	3.0%	0.414	-0.2%	1.035	3.5%	0.415	0.0%
E1 → E2	0.300	0.0%	0.740	0.3%	1.978	559.3%	0.777	5.3%
E1 → E4	0.300	0.0%	0.740	0.3%	1.976	558.7%	0.777	5.3%
Model 2B (formative indicators correlated at 0.40)								
K → E1	3.200	6.7%	0.304	0.7%	0.344	-88.5%	0.165	-45.4%
K → E3	1.030	3.0%	0.414	-0.2%	1.036	3.6%	0.415	0.0%
E1 → E2	0.301	0.3%	0.807	0.2%	1.538	412.7%	0.828	2.9%
E1 → E4	0.301	0.3%	0.807	0.2%	1.537	412.3%	0.828	2.9%
Model 2C (formative indicators correlated at 0.70)								
K → E1	3.274	9.1%	0.262	0.8%	0.220	-92.7%	0.082	-68.5%
K → E3	1.031	3.1%	0.414	-0.2%	1.036	3.6%	0.415	0.0%
E1 → E2	0.302	0.7%	0.846	0.1%	1.338	346.0%	0.843	-0.2%
E1 → E4	0.302	0.7%	0.846	0.1%	1.337	345.7%	0.843	-0.2%

Note: Unstd. = average of unstandardized paths over replications.
Std. = average of standardized paths over replications.
% Dev. = average deviation from population value, calculated as (average of paths – population value) / population value.

Table E2. Average Standardized and Unstandardized Paths, and Deviations from Population Values (Models 3 and 4, over 1,500 replications)

	Correctly Specified Models				Misspecified Models			
Path	**Unstd.**	**% Dev.**	**Std.**	**% Dev.**	**Unstd.**	**% Dev.**	**Std.**	**% Dev.**
Model 3A (formative indicators correlated at 0.10)								
K → E1	0.300	0.0%	0.711	-0.1%	1.771	490.3%	0.760	6.7%
K → E3	0.300	0.0%	0.713	0.1%	1.772	490.7%	0.760	6.7%
E1 → E2	0.601	0.2%	0.647	-0.3%	0.601	0.2%	0.647	-0.3%
E1 → E4	0.606	1.0%	0.653	0.6%	0.607	1.2%	0.654	0.8%
K → E4	-0.002		-0.005		-0.010		-0.006	
Model 3B (formative indicators correlated at 0.40)								
K → E1	0.301	0.3%	0.791	0.0%	1.456	385.3%	0.825	4.3%
K → E3	0.301	0.3%	0.792	0.1%	1.456	385.3%	0.825	4.3%
E1 → E2	0.600	0.0%	0.698	-0.3%	0.600	0.0%	0.698	-0.3%
E1 → E4	0.605	0.8%	0.704	0.6%	0.606	1.0%	0.705	0.7%
K → E4	-0.002		-0.004		-0.007		-0.005	
Model 3C (formative indicators correlated at 0.70)								
K → E1	0.302	0.7%	0.836	0.0%	1.311	337.0%	0.851	1.8%
K → E3	0.302	0.7%	0.836	0.0%	1.310	336.7%	0.852	1.9%
E1 → E2	0.600	0.0%	0.736	0.0%	0.600	0.0%	0.736	0.0%
E1 → E4	0.606	1.0%	0.742	0.8%	0.606	1.0%	0.743	1.0%
K → E4	-0.002		-0.005		-0.006		-0.005	
Model 4A (formative indicators correlated at 0.10)								
K → E1	3.185	6.2%	0.377	0.3%	0.362	-87.9%	0.268	-28.7%
K → E3	1.030	3.0%	0.414	-0.2%	1.043	4.3%	0.418	0.7%
E1 → E2	0.300	0.0%	0.740	0.3%	1.978	559.3%	0.777	5.3%
E1 → E4	0.300	0.0%	0.740	0.3%	1.976	558.7%	0.777	5.3%
E1 → E3	-0.001		-0.002		-0.009		-0.005	
Model 4B (formative indicators correlated at 0.40)								
K → E1	3.193	6.4%	0.304	0.7%	0.345	-88.5%	0.166	-45.0%
K → E3	1.030	3.0%	0.415	0.0%	1.040	4.0%	0.417	0.5%
E1 → E2	0.301	0.3%	0.807	0.2%	1.538	412.7%	0.828	2.9%
E1 → E4	0.301	0.3%	0.807	0.2%	1.537	412.3%	0.828	2.9%
E1 → E3	-0.001		-0.003		-0.005		-0.004	
Model 4C (formative indicators correlated at 0.70)								
K → E1	3.270	9.0%	0.262	0.8%	0.222	-92.6%	0.083	-68.1%
K → E3	1.031	3.1%	0.415	0.0%	1.039	3.9%	0.417	0.5%
E1 → E2	0.302	0.7%	0.846	0.1%	1.338	346.0%	0.843	-0.2%
E1 → E4	0.302	0.7%	0.846	0.1%	1.337	345.7%	0.843	-0.2%
E1 → E3	-0.001		-0.003		-0.003		-0.003	

Note: Unstd. = average of unstandardized paths over replications.
Std. = average of standardized paths over replications.
% Dev. = average deviation from population value, calculated as (average of paths – population value) / population value. Not calculated for the K → E4 path, which is zero at the population value.

Table E3. Statistical Power Results (Models 1 and 2, over 1,500 replications)

Model	K → E1	K → E3	E1 → E2	E1 → E4
Model 1A (indicators correlated at 0.10)				
Correctly specified	100.0%	100.0%	100.0%	100.0%
Misspecified	100.0%	100.0%	100.0%	100.0%
Model 1B (indicators correlated at 0.40)				
Correctly specified	100.0%	100.0%	100.0%	100.0%
Misspecified	100.0%	100.0%	100.0%	100.0%
Model 1C (indicators correlated at 0.70)				
Correctly specified	99.9%	99.9%	100.0%	100.0%
Misspecified	100.0%	100.0%	100.0%	100.0%
Model 2A (indicators correlated at 0.10)				
Correctly specified	100.0%	99.9%	100.0%	100.0%
Misspecified	75.4%	99.7%	100.0%	100.0%
Model 2B (indicators correlated at 0.40)				
Correctly specified	99.9%	99.9%	100.0%	100.0%
Misspecified	44.1%	99.7%	100.0%	100.0%
Model 2C (indicators correlated at 0.70)				
Correctly specified	98.9%	99.8%	99.5%	99.5%
Misspecified	14.9%	99.7%	100.0%	100.0%

Table E4. Statistical Power and Type I Error Results (Model 3, over 1,500 replications)

Model	K → E1	K → E3	E1 → E2	E1 → E4	K → E4
Model 3A (indicators correlated at 0.10)					
Correctly specified	100.0%	100.0%	100.0%	100.0%	4.1%
Misspecified	99.1%	100.0%	100.0%	100.0%	3.9%
Model 3B (indicators correlated at 0.40)					
Correctly specified	100.0%	100.0%	100.0%	100.0%	3.7%
Misspecified	100.0%	100.0%	100.0%	100.0%	4.3%
Model 3C (indicators correlated at 0.70)					
Correctly specified	100.0%	100.0%	99.9%	99.9%	3.0%
Misspecified	100.0%	100.0%	100.0%	100.0%	4.2%

Note: For the K → E4 path, reported values represent Type I error occurrence.

Table E5. Statistical Power and Type I Error Results (Model 4, over 1,500 replications)

Model	K → E1	K → E3	E1 → E2	E1 → E4	E1 → E3
Model 4A (indicators correlated at 0.10)					
Correctly specified	100.0%	99.6%	100.0%	100.0%	4.1%
Misspecified	73.4%	99.5%	100.0%	100.0%	3.7%
Model 4B (indicators correlated at 0.40)					
Correctly specified	99.9%	99.8%	100.0%	100.0%	4.0%
Misspecified	41.7%	99.7%	100.0%	100.0%	4.4%
Model 4C (indicators correlated at 0.70)					
Correctly specified	98.8%	99.7%	99.6%	99.6%	3.1%
Misspecified	14.3%	99.7%	100.0%	100.0%	4.9%

Note: For the E1 → E3 path, reported values represent Type I error occurrence.

References

Jarvis, C., MacKenzie, S., and Podsakoff, P. 2003. "A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research," *Journal of Consumer Research* (30), pp. 199-218.
Marsh, H., Wen, Z., and Hau, K. 2004. "Structural Equation Models of Latent Interactions: Evaluation of Alternative Estimation Strategies and Indicator Construction," *Psychological Methods* (9:3), pp. 275-3000.
Wright, S. 1934. "The Method of Path Coefficients," *Annals of Mathematical Statistics* (5:3), pp. 161-215.