On Comparing Results from CB-SEM and PLS-SEM: Five Perspectives and Five Recommendations

By Edward E. Rigdon, Marko Sarstedt and Christian M. Ringle

To estimate structural equation models, researchers can draw on two main approaches: Covariance-based structural equation modeling (CB-SEM) and partial least squares structural equation modeling (PLS-SEM). Concerns about the limitations of the different approaches might lead researchers to seek reassurance by comparing results across approaches. But should researchers expect the results from CB-SEM and PLS-SEM to agree, if the structure of the two models is otherwise the same? Differences in philosophy of science and different expectations about the research situation underlie five different perspectives on this question. We argue that the comparison of results from CB-SEM and PLS-SEM is misleading and misguided, capable of generating both false confidence and false concern. Instead of seeking confidence in the comparison of results across methods, which differ in their specific requirements, computational procedures, and imposed constraints on the model, researchers should focus on more fundamental aspects of research design. Based on our discussion, we derive recommendations for applied research using SEM.

Introduction

"...it should always be the aim of the experimenter not to revel in statistical methods (when he does revel and not swear) but steadily to diminish, by continual improvement of his experimental methods, the necessity for their use and the influence they have on his conclusions."
(Yule, 1921, p. 106)

Statistical methods abstract away from the data, which researchers can see, to things that are unknown, whether those things are population parameters like means or variances, or unobserved conceptual variables – psychological attributes like customer satisfaction. Are these unknowns being estimated or represented correctly? A researcher's reputation and the fortunes of clients ride on the soundness of results. Simulations can show the sorts of results that will occur when certain assumptions are valid and pre-defined conditions hold, but in real applications, these assumptions and background conditions hold only to a limited degree, itself unknown. Moreover, statistical models, like all models, are only approxima-

Edward E. Rigdon is Professor of Marketing in the J. Mack Robinson College of Business, Georgia State University, P. O. Box 3992, Atlanta, GA 30302-3992, USA, Phone: +1/404 413 7674, E-Mail: erigdon@gsu.edu.

Marko Sarstedt is Professor at the Department of Marketing, Otto-von-Guericke-University Magdeburg, Universitaetsplatz 2, 39106 Magdeburg, Germany, Phone: +49/391 67 58625, E-Mail: marko.sarstedt@ovgu.de.
* Corresponding author

Christian M. Ringle is Professor and Managing Director of the Institute of Human Resource Management and Organisations, Hamburg University of Technology (TUHH), Am Schwarzenberg-Campus 4, 21073 Hamburg, Germany, Phone: +49/40 42878 4419, E-Mail: c.ringle@tuhh.de.

Acknowledgements: Even though this research does not explicitly refer to the use of the statistical software SmartPLS (http://www.smartpls.com), Ringle acknowledges a financial interest in SmartPLS.

Publication	Focus of the Study
Binz Astrachan, Patel, & Wanzenried (2014)	Comparison of CB-SEM and PLS-SEM for model development and testing in family business research.
Chin (2010)	Comparison of CB-SEM and PLS-SEM while mainly offering guidelines on how to write a PLS-SEM report.
Chin & Newsted (1999)	Contrasts the efficacy of CB-SEM and PLS-SEM in small sample research.
Chin, Peterson, & Brown (2008)	General discussion of SEM in marketing, which also highlights differences between CB-SEM and PLS-SEM.
Fornell & Bookstein (1982)	Comparison of parameter estimation in CB-SEM and PLS-SEM.
Gefen, Rigdon, & Straub (2011)	Guidelines for using SEM in administrative and social science research, including aspects on how to choose between CB-SEM and PLS-SEM.
Gefen, Straub, & Boudreau (2000)	Tutorial on CB-SEM and PLS-SEM, including a general comparison.
Hair et al. (2011)	Comparison of CB-SEM and PLS-SEM and guidelines for method choice.
Hair, Sarstedt, Ringle, & Mena, (2012)	Comparison of CB-SEM and PLS-SEM, while providing a review of PLS-SEM use in marketing research.
Hair, Hollingsworth, Randolph, & Chong (2017b)	Updated and expanded assessment of PLS-SEM use in information systems research, while offering rules of thumb for choosing between CB-SEM and PLS-SEM.
Jöreskog & Wold (1982)	Explanation of methodological similarities and differences between CB-SEM and PLS-SEM.
Kaufmann & Gaeckler (2015)	Comparison of CB-SEM and PLS-SEM, while providing a review of PLS-SEM use in supply chain management research.
Lee, Petter, Fayard, & Robinson (2011)	Comparison of CB-SEM and PLS-SEM, while providing a review of PLS-SEM use in accounting research.
Lohmöller (1989)	Discusses the efficacy of CB-SEM and PLS-SEM for structural versus predictive modeling.
Peng and Lai (2012)	Comparison of CB-SEM and PLS-SEM, while providing a review of PLS-SEM use in operations management research.
Richter, Sinkovics, Ringle, & Schlägel (2016)	Review of CB-SEM and PLS-SEM use in international business research.
Sarstedt, Hair, Ringle, Thiele, & Gudergan (2016)	Contrasts the underlying assumptions of CB-SEM and PLS-SEM with regard to the nature of the measurement models and the data.
Sarstedt, Ringle, Smith, Reams, & Hair (2014)	Introduction of PLS-SEM for family business researchers, while offering a brief comparison with CB-SEM.
Scholderer & Balderjahn (2005)	Discusses methodological differences between CB-SEM and PLS-SEM along with popular misconceptions regarding the methods' use. Provides recommendations for SEM use.
Scholderer & Balderjahn (2006)	A similar focus as in Scholderer and Balderjahn (2005) but with a much stronger focus on the statistical conceptions underlying the SEM types.
Willaby, Costa, Burns, MacCann, & Roberts (2015)	Introduction of PLS-SEM for testing complex models in psychology, including a comparison with CB-SEM.

Tab. 1: Conceptual Comparisons of CB-SEM and PLS-SEM

tions of a complex reality: both their strength and their weakness lie in simplification, which involves error.

In these difficult straits, researchers naturally look for assurance, and one logical approach is to seek similarity in results from different methods. If a researcher is not sure which of several methods to use, then using multiple methods and achieving similar results seems to offer assurance that, at the least, the choice of method carries no substantial consequences. Obtaining different results, however, may suggest that one method is "correct" and the others are "wrong," or at least that conclusions are method-dependent.

Structural equation modeling (SEM), almost from its very beginning, has been divided between covariance-based SEM (CB-SEM; Jöreskog, 1967; Jöreskog, 1969, 1971) and composite-based SEM (Tenenhaus, Esposito Vinzi, Chatelin, & Lauro, 2005; Wold, 1974, 1982). One of the first composite-based approaches, partial least squares SEM (PLS-SEM), was conceived as an alternate means for accomplishing the same goal as the CB-SEM approach, with advantages in some instances, but with disadvantages in situations where the necessary conditions supporting the optimal properties of the CB-SEM approach could be expected to hold (e.g., Chin, 1998; Hair, Ringle, & Sarstedt, 2011; Jöreskog & Wold, 1982). Numerous studies have reviewed the concrete conditions that favor the use of either method, focusing on aspects such as their efficacy for estimating reflectively vs. formatively specified measurement models, distributional assumptions, and sample size requirements (Table 1). Rather than trust that these conditions hold in a particular instance, researchers typically compare results from PLS-SEM and other composite-based SEM methods such as generalized structured component analysis (GSCA; Hwang, 2009), or even regressions based on sum scores (e.g., Goodhue, Lewis, & Thompson 2012) with those from CB-SEM on the grounds of simulated data – Table 2 offers an overview of prior simulation studies comparing CB-SEM, PLS-SEM, and related estimators.

Publication	Data Generation	Model Estimation Methods	Key Findings
Dijkstra & Henseler (2015b)	Factor model	CB-SEM including the ML, GLS, WLS, DWLS, and ULS approaches, as well as PLS-SEM, PLSc, and OLS regression with sum scores	The GLS and WLS (CB-SEM) methods exhibit non-convergence issue in certain situations. PLS-SEM and regression with sum scores provide less accurate estimates, while PLSc and especially the ML (CB-SEM) methods provide consistent outcomes. Inconsistent estimates entail Type I and Type II errors.
Goodhue, Lewis, & Thompson (2012)	Factor model	ML-based CB-SEM, PLS-SEM, and OLS regression with sum scores	Both PLS-SEM and regression with sum scores are less accurate than CB-SEM. PLS-SEM performs at equal levels compared with the other techniques in terms of statistical power and avoidance of false positives. These findings also hold for small sample sizes.
Hwang, Malhotra, Kim, Tomiuk, & Hong (2010)	Factor model	ML-based CB-SEM, PLS-SEM, and GSCA	When the model is correctly specified, CB-SEM tends to better recover the parameters than PLS-SEM and GSCA (i.e., CB-SEM has a higher accuracy).
Lu, Kwan, Thomas, & Cedzynski (2011)	Factor model	CB-SEM including the Croon, Skrondal-Laake, and ML approaches, as well as PLS-SEM	The CB-SEM estimates converge to the pre-specified parameter value as the sample size increases (i.e., the estimates have a low relative bias and are consistent). PLS-SEM estimates are not consistent and show higher variability in coverage (i.e., an increase of statistical power when the sample size increases); this variability declines with more indicators per measurement models and higher outer loadings. For small sample sizes, PLS-SEM has the highest statistical power.
Reinartz, Haenlein, & Henseler (2009)	Factor model	ML-based CB-SEM and PLS-SEM	If the sample size is large (e.g., more than 250 observations), CB-SEM has higher parameter accuracy and consistency than PLS-SEM. When the sample size is small, in comparison with CB-SEM, PLS-SEM has always larger or equal statistical power.
Sarstedt et al. (2016)	Composite model / factor model	ML-based CB-SEM, PLS-SEM, and PLSc	PLS-SEM shows (almost) no bias when estimating data from a composite model population. In contrast, CBSEM and PLSc estimations show severe biases. When estimating data from common factor populations with more than 250 observations, CB-SEM's and PLSc's bias is small and diminishes with more data. PLS-SEM returns to some extent biased results. For small sample sizes, in comparison with CB-SEM and PLSc, this bias of PLS-SEM is relatively low.

Note: DWLS = diagonally weighted least squares, GLS = generalized least squares, GSCA = generalized structured component analysis, ML = maximum likelihood, OLS = ordinary least squares, PLS-SEM = partial least squares structural equation modeling, PLSc = consistent partial least squares structural equation modeling, ULS = unweighted least squares, WLS = weighted least squares.

Tab. 2: Simulation Studies Comparing Parameter Recovery Capabilities of CB-SEM and PLS-SEM[1]

These studies sparked an at times heated debate among scholars, particularly about the merits of the PLS-SEM method, which resulted in the formation of two opposing camps. One group of scholars, supportive of the (composite-based) PLS-SEM method, has emphasized the method's prediction-orientation and capabilities to handle complex models, small sample sizes, and formatively specified constructs (e.g., Chin, 2010; Hair et al., 2011; Jöreskog & Wold, 1982). The other group has noted that PLS-SEM is not a (factor-based) latent variable method, producing biased and inconsistent parameter estimates (e.g., Rönkkö, Antonakis, McIntosh, & Edwards, 2016; Rönkkö & Evermann, 2013; Rönkkö, McIntosh, & Antonakis, 2015). Naturally, such debates and particularly forceful rejections of one method such as "there is no use for PLS whatsoever" (Antonakis, Bendahan, Jacquart, & Lalive, 2010, p. 1103) or researchers should "discontinue the use of PLS" (Rönkkö et al., 2016, p. 24) might upset researchers, particularly PhD students. If PLS-SEM produces "wrong" results whereas CB-SEM produces "correct" results, why would anyone ever support the use of PLS-SEM?

But should researchers expect results from CB-SEM and PLS-SEM to agree, if the structure of the two models being estimated is otherwise the same? This question is statistical at its core, but it also goes beyond statistics, addressing the role of the model in the research enterprise. As we will argue in this manuscript, there is no univocal answer to this question. As the dominance of empiricism in philosophy of science circles has given way to realist perspectives, so have views changed regarding the nature of the conceptual variables and constructs that populate theoretical and statistical models in the social sciences – see, for example, Bollen (1989) and Bollen and Diamantopoulos (2017).

In this manuscript, we highlight five different perspectives on comparing results from CB-SEM and PLS-SEM. These perspectives imply that the universal rejection of one method over the other is shortsighted as such a step necessarily rests on assumptions about unknown entities in a model and the parameter estimation. We ar-

[1] Other simulation studies that compare CB-SEM and PLS-SEM address topics such as model fit (Dijkstra & Henseler, 2015a) and the methods' capabilities for multigroup analysis (Qureshi & Compeau, 2009), and prediction (Evermann & Tate, 2016).

gue that researchers' functional background and adherence to a specific position in philosophy of science contribute to the confusion over which method is "right" and which one is "wrong." Based on our descriptions, we offer five recommendations that share a common theme: Any empirical comparison of results from CB-SEM and PLS-SEM – despite considerable research interest (Table 2) – is misguided, capable of providing both false confidence and false concern. Instead of seeking confidence in the comparison of results from the different approaches, researchers should instead focus on more fundamental aspects of modeling, measurement, and statistical analysis.

Five perspectives

Preamble

The following five perspectives have in common that they abstract away from the concrete differences between the model estimates under specific conditions as typically researched in simulation studies. Instead, they address the underlying philosophical, theoretical, and conceptual aspects of model building and estimation to provide practical recommendations for researchers in marketing and other disciplines. We identified these five perspectives to generically systematize the broad scope of relevant issues underlying SEM. Discussing them in the proposed order does not impose a hierarchy but supports the flow of argument – all perspectives have the same raison d'être.

Perspective #1: Different Estimators

From one perspective, CB-SEM and PLS-SEM should yield the same results, because they are viewed as two methods intended to accomplish the same statistical end: estimating a series of structural equations that (1) represent causal processes, and (2) can be modeled pictorially to enable a clear conceptualization of the theory under study (Byrne, 1998). With exploratory techniques, whether factor analysis or principal components analysis, the end was data reduction – reducing the dimensionality of the data (e.g., Velicer & Jackson, 1990; Widaman, 1993). On this basis, Velicer and Jackson (1990, p. 22) argued that results from factor analysis and principal components analysis were generally comparable and that "there is no basis to assume that either method is more accurate." The confirmatory thrust which really launched SEM can be traced to Spearman's (1904) works on general intelligence and his insistence that the common factor derived in his true score model was nothing less than general intelligence itself. The true score was the lone systematic cause of variation across multiple observed variables – and, Spearman (1904) alleged, across almost any valid test of any ability whatsoever.

Factor analysis achieved a purification, boiling away the unwanted elements within the observed variables, leaving only items with high correlations to represent the underlying unobserved entity. Spearman's (1904) theory of a single general cause gave way to models with multiple correlated attributes or conceptual variables. Still, factor analysis, along with econometric and path models, inspired Jöreskog's (1969) work devising an inferential test for an a priori factor model with structured relations. And Herman Wold, Jöreskog's "Doktorvater", the originator of PLS-SEM (Wold, 1966, 1974), in turn, was inspired to find least squares approaches for doing the same thing (Dijkstra, 2014), in a technology environment where computing time was in short supply. This composite-based method was invented with the explicit intent of approximating results from Jöreskog's (1969) maximum likelihood confirmatory factor analysis. So naturally researchers have evaluated the quality of an approximation against the standard being approximated, starting with the early works of Areskoug (1982), Jöreskog and Wold (1982), and Lohmöller (1989), and continuing with numerous simulation studies contrasting CB-SEM and PLS-SEM (Table 2).

If PLS-SEM is an approximation of the CB-SEM approach, then researchers have a right to expect that its results will closely adhere to those produced by CB-SEM. If not, then the approximation is deficient and should be discarded. This, indeed, is the line taken by authors who have been especially critical of PLS-SEM, noting that its "path estimates are biased" (Westland, 2015, p. 37) or that PLS-SEM entails "biased and inconsistent estimation" (Rönkkö et al., 2016, p. 14). Calling for the abandonment of PLS-SEM seems to be the logical consequence. Potential doubts of such a strong call bring us to the second perspective.

Perspective #2: Different Models

From a second perspective, one should *not* expect the same results from CB-SEM and PLS-SEM, because the statistical models are not equivalent. Even though Herman Wold adopted the verbal and even the graphical representations of factor analysis for his composite-based alternative, and embraced the goal of producing similar results (Jöreskog & Wold, 1982), he never claimed that his method was a form of factor analysis, even when suggesting that PLS-SEM is "deliberately approximate" to CB-SEM (Hui & Wold, 1982, p. 127).

CB-SEM and PLS-SEM employ different statistical models (Jöreskog & Wold, 1982), which assume fundamentally different measurement philosophies (Sarstedt et al., 2016). CB-SEM models the constructs as common factors that explain the covariation between their associated indicators. The scores of these common factors are neither known nor needed in the estimation of model parameters. That is, common factors can generally not be expressed solely as a function of the data in the model – a part of the common factor remains an arbitrary quantity, subject to rules but capable of taking on an infinite range of values (e.g., Mulaik & McDonald, 1978; Schönemann & Wang, 1972; Steiger, 1979). In fact, only under very special conditions do factors become determinate (Krijnen, Dijkstra,

& Gill, 1998). Composites in PLS-SEM, by contrast, are determinate functions of the other variables in the model.[2] A determinate function does not mean just predicted by but rather implies that the composite is by definition equal to a (usually) weighted sum of a specific subset of the model's manifest variables. This means that, given a set of parameter estimates and a new case or set of values for the observed variables, or even a what-if involving a hypothetical set of values, a researcher – or an auditor following a researcher's trail – can calculate the implications of those observed values for the dependent variables in the model (Shmueli, Ray, Velasquez Estrada, & Chatla, 2016). Such a clear audit trail does not exist for indeterminate factors.

Using the composites as input, PLS-SEM applies a series of regressions with the objective of maximizing the explained variance of the endogenous construct(s). As with predictor variables in a regression, however, covariances among the manifest variables in PLS-SEM and other composite-based method are generally unconstrained. Even in variations, like Mode A estimation in PLS-SEM, where standard treatments depict the observed variables as dependent on the composite, there is no constraint on the residuals for those observed variables (e.g., Tenenhaus, 2008; Tenenhaus et al., 2005). By contrast, CB-SEM is not possible with fully unconstrained residuals, because the model will quickly have negative degrees of freedom (Hoyle, 2014). Wold (1985, p. 584) suggested that PLS-SEM users should check the assumption of uncorrelated residuals as "a partial test of the realism of the model." However, there is no compelling reason to do so as PLS-SEM and composite-based methods in general do not logically imply that assumption (e.g., Gefen et al., 2011). Hence, while CB-SEM and PLS-SEM may draw on the same *theoretical model*, the *statistical models* assumed by the methods differ fundamentally due to the methods' differing assumptions, requirements, and imposed constraints (McDonald, 1996; Sarstedt et al., 2016; Tenenhaus, 2008).

In light of the differences in the methods' statistical models, Marcoulides, Chin, and Saunders (2012) referred to any contrasting of CB-SEM and PLS-SEM as "comparing apples with oranges" (p. 725), a concern that has been voiced by many other authors in similar form (e.g., Hwang et al., 2010; Lohmöller, 1989; Schneeweiß, 1991). Responding to this criticism, researchers have recently started evaluating the performance of methods such as PLS-SEM and GSCA using composite model data (e.g., Becker, Rai, & Rigdon, 2013; Hair, Hult, Ringle, Sarstedt, & Thiele, 2017d). These studies show that PLS-SEM and GSCA are consistent estimators when the underlying population is composite model-based, contradicting prior research that assumed factor model-based populations, where these methods have been shown to slightly overestimate measurement model parameters and underestimate structural model parameters (Table 2).[3] Relatedly, Sarstedt et al. (2016) recently showed that CB-SEM's parameter bias can be substantial when (erroneously) used on data from composite model populations. Jointly, these results empirically substantiate that the two approaches to SEM are neither the same nor interchangeable, though transformations have been found to turn canonical correlation results into maximum likelihood estimates for the inter-battery factor model (Browne, 1979; Tucker, 1958), or to transform PLS-SEM parameter estimates into consistent estimates of factor model parameters (Dijkstra & Henseler, 2015b).

Perspective #3: Same Phenomena

From a third perspective, researchers *should* expect similar results from CB-SEM and PLS-SEM, because both methods are tapping the same real-world phenomena – unobservable conceptual variables like quality, customer satisfaction, and loyalty in marketing. Here, though, differing approaches to philosophy of science provide different views of what "real-world phenomena" means, in an SEM context. Of course, broad approaches to philosophy of science, even individual approaches, are far from monolithic, more like movements than individual positions, offering a range of variation even within a single general position. Still, philosophy of science has experienced substantial change over the years, being successively dominated by forms of idealism, empiricism, and realism, with important implications for SEM.

The classical idealist position rejected the notion of any knowable reality outside the mind (Hunt, 1991). Such a reality might exist, but there was no way to prove it. The mind is all there is, and notions like "customer satisfaction" are purely mental constructions. In reaction, empiricism insisted that the mind was not competent to define reality (Hunt, 1991). Science, it argued, was constrained by data, limited only to that which could be systematically observed. So "customer satisfaction," for example, must be defined in terms of data, a function of observed variables. Operationalism, an extreme form of empiricism, explicitly defined every variable in terms of specific data gathering operations, and denied existence to anything, including central psychological variables like intelligence, beyond that specific set of operations (Chang, 2009). Realism, by contrast, argues that a mind-independent world exists, and that science can achieve understanding about that world, even about conceptual variables like "customer satisfaction," which cannot be di-

[2] Note that factor models and composite models are not equivalent to reflective and composite indicators. While the model types indicate whether indicator covariances (factor model) or linear combinations of indicators (composite model) define the nature of the data (Sarstedt et al., 2016), the indicator type refers to the theoretical specification of the measurement models based on measurement theory (Bollen, 2011).

[3] In fact, PLS-SEM pioneers claimed only that estimates based on factor model data were "consistent at large," meaning that they correspond to those produced by factor-based SEM when sample size and number of measurement indicators are infinite, provided that the data stem from a factor model population (Hui & Wold, 1982).

rectly observed (Chakravartty, 2007; Haig & Evers, 2016). For the realist, the conceptual variable exists independent of observation and transcends data – though research on unobservable entities is fraught with challenges when it comes to validating inferences.

But if both CB-SEM and PLS-SEM are tapping the same phenomena, what phenomena are those? For the true empiricist, data are the object of study and the only legitimate subject matter of science. A term like "customer satisfaction" must refer to a function of data. When Wold (1975), for example, described his "Nonlinear Iterative Partial Least Squares" method, a precursor to PLS-SEM, he wrote about the composites in his model serving as proxies for the group of observed variables that formed each composite. Similarly, an empiricist using CB-SEM might define "customer satisfaction" as the common factor derived from a set of indicators. Consistent with his empiricist orientation, Spearman (1904) had no problem equating general intelligence itself with his statistical rendering of it. Empiricism continued to dominate philosophy of science in the social sciences throughout the period when SEM was emerging, and it has had a profound impact on thinking in the field, as it has on statistical education generally.

The tight linkage between modern psychometrics and CB-SEM depends upon an empiricist identification of the common factor with the conceptual variable. If the common factor extracted from a set of observed variables is, in fact, identical with the conceptual variable that a researcher seeks to study, then one can assess the construct validity by examining the strength of the relationship between factor and indicators (i.e., the loadings; Homburg & Giering, 1997). The disturbances of the individual items – the part of each indicator not associated with the factor – are then "measurement errors." Construct validity can be assessed purely as a function of model parameters. There is no need to ask whether a factor labeled "customer satisfaction" actually behaves like the conceptual variable "customer satisfaction," because the factor and the conceptual variable by definition are the same thing. Empiricists using PLS-SEM would grant the same status to their composites.

By contrast, the realist will recognize "customer satisfaction" as being an actual attribute of persons, a real feature of the world beyond the laboratory, which is captured imperfectly by statistical analysis. Both the factors in CB-SEM and the composites in PLS-SEM are proxies for the conceptual variables themselves. Inferences based on the use of proxy variables must face additional scrutiny, because invalid proxies will invalidate inferences based upon them. A realist cannot test construct validity purely within the confines of a statistical model, because the realist cannot assume that proxy and conceptual variable are identical. A realist's proxies refer to actual entities external to the statistical model, and so evidence for construct validity must also refer to those entities, comparing the behavior of each proxy to what is known or believed about the actual psychological attribute. The disturbances in a factor model are not "measurement errors," but only disturbances, so their absence in PLS-SEM never meant that PLS-SEM "does not compensate for measurement error" (Goodhue et al. 2012, p. 981; see also Rönkkö 2014). Instead, "measurement error" is better taken to describe the gap between proxy and conceptual variable, whether a researcher uses both CB-SEM and PLS-SEM. Still, for the realist, both approaches to SEM are attempts to study precisely the same thing, and so, in a perfect world, the different approaches absolutely must agree, or else there is a significant validity issue – though that agreement need not mean that either method's results are "correct."

Perspective #4: Imperfect World

Yet, researchers must acknowledge the practical problems that come with applying statistical methods in actual research situations. From this fourth perspective, researchers should not expect similar results from CB-SEM and PLS-SEM because, even though they ultimately address the same phenomena, the methods are typically applied in suboptimal situations, where different approaches fall short for different reasons. Within the realm of CB-SEM, for example, maximum likelihood estimation and generalized least squares estimation are asymptotically equivalent when assumptions hold, but equivalence fails when assumptions are violated (Bollen, 1989). In this regard, the only reality is the statistical reality. There is no need for an external reality. One can fabricate data under suboptimal conditions and observe predictable discrepancies between estimation methods (Henseler, Dijkstra, Sarstedt, Ringle, Diamantopoulos, Straub, Ketchen, Hair, Hult, & Calantone, 2014).

Recent research has witnessed such approaches where researchers have used highly stylized model constellations to demonstrate limitations of the PLS-SEM method, which improve our understanding of the method's performance. However, constellations such as Rönkkö and Evermann's (2013) two-construct model with a zero relationship depict boundary conditions that have very little resemblance with real-world settings, offering no grounds for abandoning the use of this or any other SEM method. Or as Henseler et al. (2014, p. 202) note: "For all methods, no matter how impressive their pedigree (maximum likelihood being no exception), one can find situations where they do not work as advertised. One can always construct a setup where a given method, any method, 'fails'."

From a realist perspective, however, there is an external reality, and researchers want to see results from CB-SEM and PLS-SEM converging upon that reality. But these methods are often applied in conditions that undermine the ability of their factors or composites to faithfully and efficiently represent the conceptual variables that populate theoretical models. Primary problems include (1) low sample size, (2) few indicators, and (3) ex post modification.

Small sample size

SEM methods are often applied to data sets that are too small. Particularly the PLS-SEM literature has long suggested that the method can be applied when sample size is very small. This tradition carries back at least to Wold (1982), who argued that PLS-SEM "worked" even when sample size was less than the number of observed variables. PLS-SEM retains its basic functionality in such conditions because the method does not estimate all model parameters simultaneously. Instead, as its name implies, it only estimates partial model structures, one equation at a time (e.g., Tenenhaus et al., 2005). Therefore, minimum sample size requirements to produce model estimates depend on the complexity of single equations, which usually is substantially less than the complexity of the overall model. Furthermore, PLS-SEM's reliance on bootstrapping for standard errors also helps to preserve basic functionality at low sample sizes. For years, the "ten times rule" has provided an informal standard (e.g., Chin, 1998; Hair et al., 2011) according to which the minimum sample size must be ten times the largest number of predictors in any equation in the system being estimated – regardless of the size of the system overall. Numerous studies have called the legitimacy of the ten times rule into question (e.g., Goodhue et al., 2012; Kock & Hadaya, 2017; Rönkkö & Evermann, 2013), including two that used composite model data in their analyses. Specifically, Becker et al. (2013) demonstrated that, when sample size is small, PLS-SEM performs poorly in terms of out-of-sample prediction (i.e., the ability to take parameter estimates, calculated from a sample, and make predictions about the larger population). More recently, Hair et al. (2017d) found that PLS-SEM produces substantially biased parameter estimates in the measurement models when sample size is small. These results suggest that caution needs to be exercised when interpreting PLS-SEM results on an item level since the biases produced in this situation potentially cast doubt on any prioritization on the grounds of indicator weights. However, Hair et al. (2017d) also found that PLS-SEM's performance in the structural model where the relative deviations for small sample sizes are comparable to those produced by CB-SEM for sample sizes of 250 to 500 when estimating factor model-based data (Reinartz et al., 2009).

Compared to PLS-SEM, small sample sizes are much more of an issue for CB-SEM, which typically produces inadmissible solutions unless several hundred observations are available. Maximum likelihood's sample size demands were a primary motivator for Wold's developing his least squares alternative (Jöreskog & Wold, 1982), but from a realist perspective, both approaches to SEM are large sample methods (Rigdon, 2016). Of course, the sample size requirements can vary even within the same statistical method, depending on the specific criteria that are most important to researchers, as Maxwell (2000), for example, describes in connection with regression. Nevertheless, one clear conclusion from Becker et al.'s (2013) simulations is that, in the absence of a large sample size, researchers who judge their results against criteria like out-of-sample prediction will do as well or better to simply sum their multiple indicators into unit-weight composites. Similarly, Hair et al. (2017d) have shown that unit-weight composites outperform indicator-weighting methods such as PLS-SEM in terms of bias and consistency when sample size is small and measurement models have many indicators. In this situation, the bias in indicator weights due to sampling variability is higher than the bias resulting from the assignment of equal weights. With many indicators per measurement model, the average weights decrease, which amplifies this effect.

Few indicators

Besides the use of small samples, the performance of SEM methods is also hampered by the use of too few observed variables per factor/composite. The mechanism underlying the effect of using few indicators is different for CB-SEM and PLS-SEM. With composites as used in PLS-SEM, the relationship is more straightforward, a function of random variance. Let C be a composite of p weighted variables x_i (i=1,...p), i.e.,

$$C = \sum_{i=1}^{p} w_i x_i , \tag{1}$$

where the w_is are the multiplier weights for multiplying each respective variable before adding it to the composite. Then, the variance of the composite C is given by

$$\sigma_C^2 = \sum_{i=1}^{p} w_i \sigma_i^2 + 2\sum_{i=1}^{p} \sum_{j=1}^{i-1} w_i w_j \sigma_{ij} , \tag{2}$$

where σ_i^2 is the variance of x_i and σ_{ij} (i=1 to p, j=1 to i, $i \neq j$) the covariance between different indicators x_i and x_j (Mulaik, 2010, equation 3.21, p. 83). That is, the variance of a sum is equal to the sum of the components' variances plus twice the sum of their covariances, each adjusted by the weights. Random variance is orthogonal, so it plays no part in the covariances. Thus, a composite can be expected to have less random variance than its components – much less as the number of independent components increases. So the composite will be more reliable and, thus, more strongly associated with any criterion, such as the real-world phenomenon that the composite represents in the statistical model (Henseler et al., 2014; Rigdon, 2012).

With CB-SEM, the relation between method performance and number of indicators is more indirect, operating through factor indeterminacy (Grice, 2001), which declines as the number of indicators per factor and the strength of their relations with the factors increase (Guttman, 1955) – the strength of relations between factors and indicators is itself a function of the degree of random variance in the observed variables (when the factor model holds; also see Guttman, 1958). For the realist who takes an actual real-world phenomenon to be the ultimate criterion of validity, factor indeterminacy has an important consequence. While the indeterminacy of a common factor does not affect parameter estimates within the same factor model, it does limit the correlation between the indeterminate common factor and any other variable outside the given factor model (Steiger, 1979). This is important because the real-world phenomenon (e.g., the actual customer satisfaction) is itself *outside* the factor model – if it could be included within the model, there would be no need for a proxy of any kind. To summarize, the validity of measurement produced by CB-SEM and PLS-SEM is likely to be limited when researchers use few indicators in the measurement models.

Ex post modification

The ex post model modification that has become an integral element of practically every SEM study (Bagozzi & Yi, 2012; Hair, Babin, & Krey, 2017a; Hair, Hult, Ringle, & Sarstedt, 2017c; Rigdon, 2013) widens the gap between the two methods. Now, if users were aware of the real-world criterion, and could match their work-in-progress against that criterion, then CB-SEM and PLS-SEM might be observed to *converge* over the course of a research program, both approaching a better match to the criterion. Instead, researchers are invited to optimize on purely statistical criteria that, at best, are indifferent to outside matters like the real world. For CB-SEM users, the relevant standards are the χ^2 statistic or alternative fit indices (e.g., Bagozzi & Yi, 2012). For users of PLS-SEM methods, typical standards are the structural model R^2, the statistical significance of parameter estimates (e.g., Hair, Sarstedt, Pieper, & Ringle, 2012a; Hair, et al., 2012b; Ringle, Sarstedt, & Straub, 2012), and, more recently, the model's out-of-sample predictive power (Shmueli et al., 2016). In either event, these statistical standards often drive researchers into modifying their models in various ways, adding or deleting parameters and, often enough, discarding indicators or entire constructs (e.g., Cliff, 1983; Sarstedt, Ringle, Henseler, & Hair, 2014). The movement to satisfy distinct statistical criteria seems likely to drive apart CB-SEM results from PLS-SEM results. In regard to the external criterion of an actual real-world phenomenon, deleting indicators or constructs tends to increase the influence of factor indeterminacy or random variance, weakening the relationship with the external criterion (Diamantopoulos, Sarstedt, Fuchs, Wilczynski, & Kaiser, 2012).

More generally, it is implausible that ex post modifications intended to improve purely internal statistical criteria will at the same time drive results to converge upon any external criterion. Fit indices only aim at affirming attributes of the statistical model, not at affirming identity of the conceptual variables or their behavior.

Perspective #5: The Long View

The fifth and final perspective is that CB-SEM and PLS-SEM methods *should* converge on the same answers about relations between unobserved conceptual variables *if* both types of methods are applied under favorable conditions. The latter may include large samples, many indicators per measurement model, all closely tied to the phenomena under study, and a commitment across a cumulative research program to develop measures that faithfully represent particular conceptual variables. Under those conditions, it must be true that multiple approaches lead to the same conclusions, or else there is no ground for inferring validity.

Researchers, however, must address the real-world phenomena under study, and the real-world consequences that follow. Hence, they need to maintain a focus on the intended outcome rather than focusing strictly on mere statistical byproducts of a certain method. In fact, applied researchers often are not interested in exact coefficient values and precise significances but rather in finding salient relationships that warrant managerial attention. Simple graphical inspections of the data often suffice for qualitatively interpreting outcomes and substantiating a priori expectations. Statistical methods that quantify these outcomes may be taken to provide additional assurance but may also convey a spurious accuracy, given that the exact results reported depend on the many idiosyncrasies of the data, definitions, and measurement operationalizations.

Only repeated examinations of a certain phenomenon provide confidence for the prevalence of some effect (Begley & Ioannidis, 2015; Jasny, Chin, Chong, & Vignieri, 2011; Open_Science_Collaboration & Kappes, 2014). From this view, differences between CB-SEM and PLS-SEM, which typically occur at the second decimal place when the data stem from a factor model population (e.g., Reinartz et al., 2009; Sarstedt et al., 2016), are of little relevance.

Five recommendations

In most papers like this one, the "Recommendations" section, following the back-and-forth at the heart of the paper, amounts to a choice of one method or the other. However, as our prior discussions have shown, there is not the single best option as the choice of a method depends on the explicit or implicit assumptions a researcher makes regarding the phenomena under study and our ability to measure them comprehensively. Researchers in psychometrics will likely view the factor model as the standard of comparison and reject any effort to represent unobservable phenomena with composite variables as an "indicator weighting cannot meaningfully reduce the effect of measurement error in the composites" (Rönkkö et al., 2016, p. 12). Researchers in econometrics, however, will likely subscribe to the proxy nature of measurement, as their field has long dealt with issues related to the use of a proxy when a variable specified in a theoretical model was unobservable and thus unavailable for statistical model estimation (e.g., Wickens, 1972). Yet others may acknowledge the practical problems when applying statistical methods, which limit their ability to fully implement their research program. For example, researchers wishing to estimate very complex models quickly reach limits when using CB-SEM, which might motivate them to subscribe to a more pragmatic view about measuring the phenomena of interest.

Following Yule (1921), we believe that the choice between CB-SEM and PLS-SEM may be of secondary importance compared to the many research design choices that researchers face when attempting to learn about the behavior of unobserved conceptual variables. It is all too easy to be distracted by purely statistical concerns, and to fall into a sort of ritualistic behavior (Gigerenzer, 2004), engaging in the "cargo cult science" warned against by Feynman (1974). Cookbook, mechanistic approaches to research seem perennially popular. SEM's affection for overall fit indices and "cut-off values" are all part of a grand ritual aimed at gaining publication, but which may have nothing to do with learning about any real-world phenomenon (Cliff, 1983). Studies that dramatize differences in the methods (Rönkkö et al., 2016; Rönkkö & Evermann, 2013), thereby inflating the relevance of method choice, have contributed to this development. That is our first recommendation: Focus on the actual phenomenon you are supposedly studying, and don't let the modeling get in the way of the learning.

The second recommendation, also inspired by Yule (1921), is to put the other design aspects of research back on the table. Researchers can wreck a project well before the statistical analysis, but in research reports employing SEM, too often the data seem to be taken as given, as if handed down on tablets. Of course, researchers may discard individual cases or whole observed variables in the course of their statistical analysis, but that is as an alternative to critiquing the source of the data, and perhaps acknowledging that the data are entirely unsuitable for the purpose. As Sir Ronald Fisher (1938, p. 17) noted:

"Immensely laborious calculations on inferior data may increase the yield [of the information contained in the data] from 95 to 100 per cent. A gain of 5 per cent, of perhaps a small total. A competent overhauling of the process of collection, or of the experimental design, may often increase the yield ten or twelve fold, for the same cost in time and labor. To consult the statistician after the experiment is finished is often merely to ask him to conduct a *post mortem* examination. He can perhaps say what the experiment died of."

As researchers have warned before, design fundamentals are sometimes forgotten when employing high-powered statistical tools (e.g., Cliff, 1983; de Leeuw, 1985). Different aspects of the research process contribute differently to the process' ultimate success. Some research design choices such as questionnaire design contribute to putting information into the data, while others, like the choice of statistical method, contribute to extracting the information from the data. Different statistical methods may have marginal advantages in extracting information, under different conditions, but no statistical method can find information that is not there. Hence, we recommend that researchers should more strongly focus on building data sets that can credibly lead to insight.

Our third recommendation is that researchers use a technique that is consistent with the type of model that they intend to estimate – in other words, that they correctly estimate their chosen model. There has been a tendency in the literature to treat CB-SEM and PLS-SEM as if they were estimating the same model. They are not. Researchers who intend to estimate a factor model should use CB-SEM, while researchers intending to estimate a model of composites should use a composite-based method like PLS-SEM, GSCA, or another in that class (e.g., Dijkstra & Henseler, 2011; Tenenhaus & Tenenhaus, 2011). Researchers who wish to estimate a model that includes a mix of factors and composites face limited choices, the one possibility currently being Dijkstra and Henseler's (2015a, 2015b) consistent PLS method, which involves estimating a composite model and then converting (some) parameter estimates to those consistent with a factor model.

It is not always easy to have clarity on this seemingly straightforward matter, because these methods are often described with unclear language. Researchers refer to "constructs" or "latent variables" (Michell, 2013) rather than referring to either common factors or composites. "Latent variables" are a feature of conceptual or theoretical models, not statistical models. Different statistical packages can estimate either factor models or composite models, with consistent PLS having some ability to bridge the gap, and researchers should use a package that can estimate the model in which the researcher is interested.

But this leaves a central question – perhaps the most central question for the journal's special issue – still unanswered: which type of model should researchers intend to estimate? This question must be answered within a specific context. Obviously, if a researcher knows the data generating process, then they should use that same model. Unfortunately, we only know this in situations where data are fabricated, as in simulation studies. In the normal course of applied research, collected data are the result of an unknown (and probably messy) process, neutralizing any argument about matching the model to the process. Still, researchers may be participating in an established research program. Ideally, a research program progressing incrementally will yield lessons about which approach produces the best proxies – proxies that best replicate the behavior of the conceptual variables that the proxies represent. In such a case, of course, we encourage researchers to take the fullest advantage of this prior knowledge.

But then, what about research efforts that are new, and not part of an ongoing research program? Although no one would recommend this situation to any novice researcher, it probably occurs regularly. Again, consider the situation. The researcher has only a limited grasp of the conceptual variables themselves, and may have limited experience with the instrument used to collect the data. Perhaps the data are already in hand, or perhaps the data are still to be collected, leaving open the possibility that the researcher could make adjustments depending on the analytical method to be chosen.

This, indeed, is a situation for which Wold long ago recommended PLS-SEM (e.g., Wold, 1974, 1980, 1985) and which we adhere to in our fourth recommendation. Instead of fretting about conforming to the constraints of a factor model, which no one expects to hold exactly, anyway (e.g., Asparouhov, Muthén, & Morin, 2015; Cudeck & Henly, 1991; Jöreskog, 1969), in a situation that is "data-rich and theory-skeletal" (Lohmöller & Wold, 1980, p. 1), PLS-SEM enables the researcher to examine the data and evaluate many different configurations. This characteristic is particularly beneficial in big data applications, which typically focus on prediction, rely on complex models with little theoretical substantiation (Stieglitz, Dang-Xuan, Bruns, & Neuberger, 2014), and often lack a comprehensive support on the grounds of measurement theory (Rigdon, 2013).

Some purists will object that this sounds like a fishing expedition, as opposed to the theory-first single line of inference approach so often described in journal articles (and so rarely followed in practice). But an exploratory approach is not in itself deficient. Rather, the problem lies in pairing exploratory modeling with inferences that presume a single test of a clearly stated hypothesis. Statistics designed for the single hypothesis test, like the typical *p*-value, are clearly inappropriate in an exploratory environment (Wasserstein & Lazar, 2016). The products of the exploratory effort are not conclusive answers but rather clues, which may (or may not) point the diligent researcher in a profitable direction.

In recommending PLS-SEM for exploratory settings, we are not, by way of contrast, recommending the sole reliance on CB-SEM for confirmatory settings, nor are we endorsing the range of conventional notions of when to favor one or the other approach as laid out in dozens of articles (e.g., Hair et al., 2011; Kaufmann & Gaeckler, 2015; Peng & Lai, 2012). As noted previously, the choice between CB-SEM and PLS-SEM has been overemphasized in the methodological literature and is of secondary importance in a research endeavor. Importantly, this notion is not only based on a difference in quality of results – which is usually marginal in situations commonly encountered in applied research (e.g., Henseler et al., 2014) – but rather on the ease of obtaining results, which can only be suggestive of potential relationships. The supporting assumptions and constraints of the CB-SEM approach are particularly unlikely to hold in an exploratory context (Asparouhov & Muthén, 2009; Kline, 2016). Hence, the explorer's time would be better invested in exploring possibilities, rather than in chasing improbabilities.

There is never one unique model that characterizes the empirical evidence within a theoretical framework. Variations may offer theoretically justified alternatives for explaining the phenomenon under study. Therefore, our fifth recommendation is that researchers should more routinely explore theoretically justified alternative models for explaining the phenomenon under study. Model comparisons are crucial for advancing scientific knowledge, and are imperative to assess the strength of one theory over another (Canham, Cole, & Lauenroth, 2003). For example, Popper (1959) argued that comparing alternative explanations (or possible causes) is a crucial step prior to any attempt at the falsification of a theory. Similarly, Platt (1964, p. 350) argued that researchers should entertain a "conflict between ideas" by devising and comparing alternative explanations (models) in any given study. Both CB-SEM and PLS-SEM are well equipped for such comparisons. In CB-SEM, model comparisons typically rely on χ^2 difference tests when models are nested (Anderson & Gerbing, 1988) and model selection criteria when models are non-nested (Rust, Lee, & Valente, 1995). The latter set of metrics can also be used in PLS-SEM. For example, Sharma et al. (2015) compare the performance of several model selection criteria in choosing the best model in a set of competing models in a PLS-SEM context.

Conclusion

It is important to underline our first two recommendations. We believe that quality data lead to quality inferences and that no statistical method turns bad data into good data. We are also concerned that our limited recommendation of PLS-SEM – which one would easily expect given our publication track record in the field – will be misunderstood. Often enough, an author writes (as we ourselves have done, now and then), "Here is a situation where you should use method X," and all that gets quoted is, "...use method X." In response to the hostility coming from some quarters of the CB-SEM community, it is all too tempting to return fire and make broad claims for PLS-SEM. We think it is best to leave such sniping to thin-skinned politicians and encourage researchers to consider the entirety of the research process.

References

Anderson, J. C., & Gerbing, D. W. (1988). Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. *Psychological Bulletin, 103*(3), 411–423.

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On Making Causal Claims: A Review and Recommendations. *The Leadership Quarterly, 21*(6), 1086–1120.

Areskoug, B. (1982). The First Canonical Correlation: Theoretical PLS analysis and Simulation Experiments. In H. O. A. Wold & K. G. Jöreskog (Eds.), *Systems under indirect observation: Causality, structure, prediction* (Vol. 2). Amsterdam: North-Holland.

Asparouhov, T., & Muthén, B. (2009). Exploratory Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 16*(3), 397–438.

Asparouhov, T., Muthén, B., & Morin, A. J. S. (2015). Bayesian Structural Equation Modeling With Cross-Loadings and Residual Covariances. *Journal of Management, 41*(6), 1561–1577.

Bagozzi, R. P., & Yi, Y. (2012). Specification, Evaluation, and Interpretation of Structural Equation Models. *Journal of the Academy of Marketing Science, 40*(1), 8–34.

Becker, J.-M., Rai, A., & Rigdon, E. E. (2013). Predictive Validity and Formative Measurement in Structural Equation Modeling: Embracing Practical Relevance. *2013 Proceedings of the International Conference on Information Systems*, Milan.

Begley, C. G., & Ioannidis, J. P. A. (2015). Reproducibility in Science: Improving the Standard for Basic and Preclinical Research. *Circulation Research, 116* (January), 116–126.

Binz Astrachan, C., Patel, V. K., & Wanzenried, G. (2014). A Comparative Study of CB-SEM and PLS-SEM for Theory Development in Family Firm Research. *Journal of Family Business Strategy, 5*(1), 116–128.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York: Wiley.

Bollen, K. A. (2011). Evaluating Effect, Composite, and Causal Indicators in Structural Equation Models. *MIS Quarterly, 35*(2), 359–372.

Bollen, K. A., & Diamantopoulos, A. (2017). In Defense of Causal-Formative Indicators: A Minority Report. *Psychological Methods*, forthcoming.

Browne, M. W. (1979). The Maximum-Likelihood Solution in Inter-Battery Factor Analysis. *British Journal of Mathematical and Statistical Psychology, 32*(1), 75–86.

Byrne, B. M. (1998). *Structural Equation Modeling with LISREL, PRELIS, and SIMPLIS: Basic Concepts, Applications, and Programming*. Mahwah, NJ: Lawrence Erlbaum.

Canham, C. D., Cole, J. J., & Lauenroth, W. K. (2003). Models in Ecosystem Science. In C. D. Canham, J. J. Cole & W. K. Lauenroth (Eds.), *Models in Ecosystem Science* (pp. 1–10). Princton, NJ: Princton University Press.

Chakravartty, A. (2007). *A Metaphysics for Scientific Realism: Knowing the Unobservable*. Cambridge: Cambridge University Press.

Chang, H. (2009). Operationalism In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy.* Retrieved from http://plato.stanford.edu/archives/fall2009/entries/operationalism/: Stanford University.

Chin, W. W. (1998). The Partial Least Squares Approach to Structural Equation Modeling. In G. A. Marcoulides (Ed.), *Modern Methods for Business Research* (pp. 295–358). Mahwah: Erlbaum.

Chin, W. W. (2010). How to Write Up and Report PLS Analyses. In V. Esposito Vinzi, W. W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of Partial Least Squares: Concepts, Methods and Applications (Springer Handbooks of Computational Statistics Series, vol. II)* (pp. 655–690). Heidelberg, Dordrecht, London, New York: Springer.

Chin, W. W., & Newsted, P. R. (1999). Structural Equation Modeling Analysis with Small Samples Using Partial Least Squares. In R. H. Hoyle (Ed.), *Statistical Strategies for Small Sample Research* (pp. 307–341). Thousand Oaks: Sage.

Chin, W. W., Peterson, R. A., & Brown, S. P. (2008). Structural Equation Modeling in Marketing: Some Practical Reminders. *Journal of Marketing Theory and Practice, 16*(4), 287–298.

Cliff, N. (1983). Some Cautions Concerning the Application of Causal Modeling Methods. *Multivariate Behavioral Research, 18*(1), 115–126.

Cudeck, R., & Henly, S. J. (1991). Model Selection in Covariance Structures Analysis and the "Problem" of Sample Size: A Clarification. *Psychological Bulletin, 109*(3), 512–519.

de Leeuw, J. (1985). Reviews. *Psychometrika, 50*(3), 371–375.

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for Choosing Between Multi-item and Single-item Scales for Construct Measurement: A Predictive Validity Perspective. *Journal of the Academy of Marketing Science, 40*(3), 434–449.

Dijkstra, T. K. (2014). PLS' Janus Face – Response to Professor Rigdon's 'Rethinking Partial Least Squares Modeling: In Praise of Simple Methods'. *Long Range Planning, 47*(3), 146–153.

Dijkstra, T. K., & Henseler, J. (2011). Linear Indices in Nonlinear Structural Equation Models: Best Fitting Proper Indices and Other Composites. *Quality & Quantity, 45*(6), 1505–1518.

Dijkstra, T. K., & Henseler, J. (2015a). Consistent and Asymptotically Normal PLS Estimators for Linear Structural Equations. *Computational Statistics & Data Analysis, 81*(1), 10–23.

Dijkstra, T. K., & Henseler, J. (2015b). Consistent Partial Least Squares Path Modeling. *MIS Quarterly, 39*(2), 297–316.

Evermann, J., & Tate, M. (2016). Assessing the Predictive Performance of Structural Equation Model Estimators. *Journal of Business Research, 69*(10), 4565–4582.

Feynman, R. P. (1974). Cargo Cult Science: Some Remarks on Science, Pseudoscience, and Learning How Not to Fool Yourself. In C. A. C. I. o. Technology (Ed.), http://calteches.library.caltech.edu/51/2/CargoCult.htm. Pasadena, CA: California Institute of Technology.

Fisher, R. A. (1938). Presidential Address. *Sankhya: The Indian Journal of Statistics, 4*(1), 14–17.

Fornell, C. G., & Bookstein, F. L. (1982). Two Structural Equation Models: LISREL and PLS Applied to Consumer Exit-Voice Theory. *Journal of Marketing Research, 19*(4), 440–452.

Gefen, D., Rigdon, E. E., & Straub, D. W. (2011). Editor's Comment: An Update and Extension to SEM Guidelines for Administrative and Social Science Research. *MIS Quarterly, 35*(2), iii-xiv.

Gefen, D., Straub, D. W., & Boudreau, M.-C. (2000). Structural Equation Modeling Techniques and Regression: Guidelines for Research Practice. *Communications of the AIS, 4*, 1–78.

Gigerenzer, G. (2004). Mindless Statistics. *The Journal of Socio-Economics, 33*(5), 587–606.

Goodhue, D. L., Lewis, W., & Thompson, R. (2012). Does PLS Have Advantages for Small Sample Size or Non-Normal Data? *MIS Quarterly, 36*(3), 981–1001.

Grice, J. W. (2001). Computing and Evaluating Factor Scores. *Psychol Methods, 6*(4), 430–450.

Guttman, L. (1955). The Determinacy of Factor Score Matrices with Implications for Five Other Basic Problems of Common-Factor Theory. *British Journal of Statistical Psychology, 8*(2), 65–81.

Guttman, L. (1958). What Lies Ahead for Factor Analysis? *Educational and Psychological Measurement, 18*(3), 497–515.

Haig, B. D., & Evers, C. W. (2016). *Realist Inquiry in Social Science*. London: Sage.

Hair, J. F., Babin, B. J., & Krey, N. (2017a). Covariance-Based Structural Equation Modeling in the Journal of Advertising: Review and Recommendations. *Journal of Advertising, 46*(1), 163–177.

Hair, J. F., Hollingsworth, C. L., Randolph, A. B., & Chong, A. Y. L. (2017b). An Updated and Expanded Assessment of PLS-SEM in Information Systems Research. *Industrial Management & Data Systems, 117*(3), 442–458.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017c). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (2 ed.). Thousand Oaks, CA: Sage.

Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., & Thiele, K. O. (2017d). Mirror, Mirror on the Wall: A Comparative Evaluation of Composite-based Structural Equation Modeling Methods *Journal of the Academy of Marketing Science*, forthcoming.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a Silver Bullet. *Journal of Marketing Theory and Practice, 19*(2), 139–151.

Hair, J. F., Sarstedt, M., Pieper, T. M., & Ringle, C. M. (2012a). The Use of Partial Least Squares Structural Equation Modeling in Strategic Management Research: A Review of Past Practices and Recommendations for Future Applications. *Long Range Planning, 45*(5–6), 320–340.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012b). An Assessment of the Use of Partial Least Squares Structural Equation Modeling in Marketing Research. *Journal of the Academy of Marketing Science, 40*(3), 414–433.

Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., Straub, D. W., Ketchen, D. J., Hair, J. F., Hult, G. T. M., & Calantone, R. J. (2014). Common Beliefs and Reality about Partial Least Squares: Comments on Rönkkö & Evermann (2013). *Organizational Research Methods, 17*(2), 182–209.

Homburg, C., & Giering, A. (1997). Konzeptualisierung und Operationalisierung komplexer Konstrukte: ein Leitfaden für die Marketingforschung. *Marketing ZFP, 18*(1), 5–24.

Hoyle, R. H. (Ed.). (2014). *Handbook of Structural Equation Modeling*. New York, NY: Guilford.

Hui, B. S., & Wold, H. O. A. (1982). Consistency and Consistency at Large of Partial Least Squares Estimates. In K. G. Jöreskog & H. O. A. Wold (Eds.), *Systems Under Indirect Observation, Part II* (pp. 119–130). Amsterdam: North Holland.

Hunt, S. D. (1991). *Modern Marketing Theory: Critical Issues in the Philosophy of Marketing Science*. Cincinnati, OH: South-Western.

Hwang, H. (2009). Regularized Generalized Structured Component Analysis. *Psychometrika, 74*(3), 517–530.

Hwang, H., Malhotra, N. K., Kim, Y., Tomiuk, M. A., & Hong, S. (2010). A Comparative Study on Parameter Recovery of Three Approaches to Structural Equation Modeling. *Journal of Marketing Research, 47*(4), 699–712.

Jasny, B. R., Chin, G., Chong, L., & Vignieri, S. (2011). Again, and Again, and Again .... *Science, 334*(6060), 1225–1225.

Jöreskog, K. G. (1967). Some Contributions to Maximum Likelihood Factor Analysis. *Psychometrika, 32*(4), 443–482.

Jöreskog, K. G. (1969). A General Approach to Confirmatory Maximum Likelihood Factor Analysis. *Psychometrika, 34*(2), 183–202.

Jöreskog, K. G. (1971). Simultaneous Factor Analysis in Several Populations. *Psychometrika, 36*(4), 409–426.

Jöreskog, K. G. (1973). A General Method for Estimating a Linear Structural Equation System. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural Equation Models in the Social Sciences* (pp. 255–284). New York: Seminar Press.

Jöreskog, K. G., & Wold, H. O. A. (1982). The ML and PLS Techniques for Modeling with Latent Variables: Historical and Comparative Aspects. In H. O. A. Wold & K. G. Jöreskog (Eds.), *Systems Under Indirect Observation, Part I* (pp. 263–270). Amsterdam: North-Holland.

Kaufmann, L., & Gaeckler, J. (2015). A Structured Review of Partial Least Squares in Supply Chain Management Research. *Journal of Purchasing and Supply Management, 21*(4), 259–272.

Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (4 ed.). New York, NY: Guilford Press.

Kock, N., & Hadaya, P. (2017). Minimum Sample Size Estimation in PLS-SEM: The Inverse Square Root and Gamma-Exponential Methods. *Information Systems Journal*, forthcoming.

Krijnen, W., Dijkstra, T. K., & Gill, R. (1998). Conditions for Factor (In) Determinacy in Factor Analysis. *Psychometrika, 63*(4), 359–367.

Lee, L., Petter, S., Fayard, D., & Robinson, S. (2011). On the Use of Partial Least Squares Path Modeling in Accounting Research. *International Journal of Accounting Information Systems, 12*(4), 305–328.

Lohmöller, J.-B. (1989). *Latent Variable Path Modeling with Partial Least Squares*. Heidelberg: Physica.

Lohmöller, J.-B., & Wold, H. O. A. (1980). *Three-mode Path Models with Latent Variables and Partial Least Squares (PLS) Parameter Estimation*. Paper presented at the European Meeting of the Psychometric Society, Groningen, Netherlands.

Lu, I. R. R., Kwan, E., Thomas, D. R., & Cedzynski, M. (2011). Two New Methods for Estimating Structural Equation Models: An Illustration and a Comparison with Two Established Methods. *International Journal of Research in Marketing, 28*(3), 258–268.

Marcoulides, G. A., Chin, W. W., & Saunders, C. (2012). When Imprecise Statistical Statements Become Problematic: A Response to Goodhue, Lewis, and Thompson. *MIS Quarterly, 36*(3), 717–728.

Maxwell, S. E. (2000). Sample Size and Multiple Regression. *Psychological Methods, 5*(4), 434–458.

McDonald, R. P. (1996). Path Analysis with Composite Variables. *Multivariate Behavioral Research, 31*(2), 239–270.

Michell, J. (2013). Constructs, Inferences and Mental Measurement. *New Ideas in Psychology, 31*(1), 13–21.

Mulaik, S. A. (2010). *The Foundations of Factor Analysis* (2 ed.). Boca Raton, FL: Chapman & Hall/CRC Press.

Mulaik, S. A., & McDonald, R. P. (1978). The Effect of Additional Variables on Factor Indeterminacy in Models With a Single Common Factor. *Psychometrika, 43*(2), 177–192.

Open_Science_Collaboration, & Kappes, H. B. (2014). The Reproducibility Project: A Model of Large-Scale Collaboration for Empirical Research on Reproducibility. In V. Stodden, F. Leisch & R. D. Peng (Eds.), *Implementing Reproducible Research* (pp. 299–324). New York: CRC Press.

Peng, D. X., & Lai, F. (2012). Using Partial Least Squares in Operations Management Research: A Practical Guideline and Summary of Past Research. *Journal of Operations Management, 30*(6), 467–480.

Platt, J. R. (1964). Strong Inference. *Science, 146*(3642), 347–353.

Popper, K. R. (1959). *The Logic of Scientific Discovery*. London and New York: Rouledge.

Qureshi, I., & Compeau, D. R. (2009). Assessing Between-Group Differences in Information Systems Research: A Comparision of Covariance- and Component-Based SEM. *MIS Quarterly, 33*(1), 197–214.

Reinartz, W. J., Haenlein, M., & Henseler, J. (2009). An Empirical Comparison of the Efficacy of Covariance-Based and Variance-Based SEM. *International Journal of Research in Marketing, 26*(4), 332–344.

Richter, N. F., Sinkovics, R. R., Ringle, C. M., & Schlägel, C. (2016). A Critical Look at the Use of SEM in International Business Research *International Marketing Review, 33*(3), 376–404.

Rigdon, E. E. (2012). Rethinking Partial Least Squares Path Modeling: In Praise of Simple Methods. *Long Range Planning, 45*(5–6), 341–358.

Rigdon, E. E. (2013). Partial Least Squares Path Modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural Equation Modeling. A Second Course* (2 ed., pp. 81–116). Charlotte NC: Information Age Publishing.

Rigdon, E. E. (2016). Choosing PLS Path Modeling as Analytical Method in European Management Research: A Realist Perspective. *European Management Journal, 34*(6), 598–605.

Ringle, C. M., Sarstedt, M., & Straub, D. W. (2012). A Critical Look at the Use of PLS-SEM in MIS Quarterly. *MIS Quarterly, 36*(1), iii-xiv.

Rönkkö, M. (2014). The Effects of Chance Correlations on Partial Least Squares Path Modeling. *Organizational Research Methods, 17*(2), 164–181.

Rönkkö, M., Antonakis, J., McIntosh, C. N., & Edwards, J. R. (2016). Partial Least Squares Path Modeling: Time for Some Serious Second Thoughts. *Journal of Operations Management, 47–48*(November), 9–27.

Rönkkö, M., & Evermann, J. (2013). A Critical Examination of Common Beliefs About Partial Least Squares Path Modeling. *Organizational Research Methods, 16*(3), 425–448.

Rönkkö, M., McIntosh, C. N., & Antonakis, J. (2015). On the Adoption of Partial Least Squares in Psychological Research: Caveat Emptor. *Personality and Individual Differences, 87*, 76–84.

Rust, R. T., Lee, C., & Valente, E. (1995). Comparing Covariance Structure Models: A General Methodology. *International Journal of Research in Marketing, 12*(4), 279–291.

Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., & Gudergan, S. P. (2016). Estimation Issues with PLS and CBSEM: Where the Bias Lies! *Journal of Business Research, 69*(10), 3998–4010.

Sarstedt, M., Ringle, C. M., Henseler, J., & Hair, J. F. (2014). On the Emancipation of PLS-SEM: A Commentary on Rigdon (2012). *Long Range Planning, 47*(3), 154–160.

Sarstedt, M., Ringle, C. M., Smith, D., Reams, R., & Hair, J. F. (2014). Partial Least Squares Structural Equation Modeling (PLS-SEM): A Useful Tool for Family Business Researchers. *Journal of Family Business Strategy, 5*(1), 105–115.

Schneeweiß, H. (1991). Models with latent variables: LISREL versus PLS. *Statistica Neerlandica, 45*(2), 145–157.

Scholderer, J., & Balderjahn, I. (2005). PLS versus LISREL: Ein Methodenvergleich. In F. W. Bilemel, A. Eggert, G. Fassott & J. Henseler (Eds.), *Handbuch PLS-Pfadmodellierung: Methoden, Anwendung, Praxisbeispiele* (pp. 87–98). Stuttgart: Schäffer-Poeschel.

Scholderer, J., & Balderjahn, I. (2006). Was unterscheidet harte und weiche Strukturgleichungsmodelle nun wirklich? Ein Klärungsversuch zur LISREL-PLS-Frage. *Marketing ZFP, 28*(1), 57–70.

Schönemann, P. H., & Wang, M.-M. (1972). Some New Results on Factor Indeterminacy. *Psychometrika, 37*(1), 61–91.

Sharma, P. N., Sarstedt, M., Shmueli, G., & Kim, K. H. (2015). Predictive Model Selection in Partial Least Squares Path Modeling (PLS-PM). In: 2nd International Symposium on Partial Least Squares Path Modeling – The Conference for PLS Users, Seville, Spain.

Shmueli, G., Ray, S., Velasquez Estrada, J. M., & Chatla, S. B. (2016). The Elephant in the Room: Evaluating the Predictive Performance of PLS Models. *Journal of Business Research, 69*(10), 4552–4564.

Spearman, C. (1904). General Intelligence: Objectively Determined and Measured. *American Journal of Psychology, 15*(2), 201–293.

Steiger, J. H. (1979). The Relationship Between External Variables and Common Factors. *Psychometrika, 44*(1), 93–97.

Stieglitz, S., Dang-Xuan, L., Bruns, A., & Neuberger, C. (2014). Social Media Analytics. An Interduisciplinary Approach and its Implications for Information Systems. *Business & Information Systems Engineering, 56*(2), 101–109.

Tenenhaus, A., & Tenenhaus, M. (2011). Regularized Generalized Canonical Correlation Analysis. *Psychometrika, 76*(2), 257–284.

Tenenhaus, M. (2008). Component-based Structural Equation Modelling. *Total Quality Management & Business Excellence, 19*(7–8), 871–886.

Tenenhaus, M., Esposito Vinzi, V., Chatelin, Y.-M., & Lauro, C. (2005). PLS Path Modeling. *Computational Statistics & Data Analysis, 48*(1), 159–205.

Tucker, L. R. (1958). An Inter-Battery Method of Factor Analysis. *Psychometrika, 23*(2), 111–136.

Velicer, W. F., & Jackson, D. N. (1990). Component Analysis versus Common Factor Analysis: Some Issues in Selecting an Appropriate Procedure. *Multivariate Behavioral Research, 25*(1), 1–28.

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's Statement on p-Values: Context, Process, and Purpose. *The American Statistician, 70*(2), 129–133.

Westland, J. C. (2015). Partial Least Squares Path Analysis *Structural Equation Models: From Paths to Networks* (pp. 23–46). Cham: Springer International Publishing.

Wickens, M. R. (1972). A Note on the Use of Proxy Variables. *Econometrica, 40*(4), 759–761.

Widaman, K. F. (1993). Common Factor Analysis Versus Principal Component Analysis: Differential Bias in Representing Model Parameters? *Multivariate Behavioral Research, 28*(3), 263–311.

Willaby, H. W., Costa, D. S. J., Burns, B. D., MacCann, C., & Roberts, R. D. (2015). Testing Complex Models With Small Sample Sizes: A Historical Overview and Empirical Demonstration of What Partial Least Squares (PLS) Can Offer Differential Psychology. *Personality and Individual Differences, 84*, 73–78.

Wold, H. O. A. (1966). Estimation of Principal Components and Related Models by Iterative Least Squares. In P. R. Krishnaiah (Ed.), *Multivariate Analysis* (pp. 391–420). New Youk: Academic Press.

Wold, H. O. A. (1974). Causal Flows with Latent Variables: Partings of Ways in the Light of NIPALS Modelling. *European Economic Review, 5*(1), 67–86.

Wold, H. O. A. (1975). Path Models with Latent Variables: The NIPALS Approach. In H. M. Blalock, A. Aganbegian, F. M. Borodkin, R. Boudon & V. Capecchi (Eds.), *Quantitative Sociology: International Perspectives on Mathematical and Statistical Modeling* (pp. 307–357). New York et al.: Academic Press.

Wold, H. O. A. (1980). Model Construction and Evaluation when Theoretical Knowledge is Scarce: Theory and Application of PLS. In J. Kmenta & J. B. Ramsey (Eds.), *Evaluation of Econometric Models* (pp. 47–74). New York: Academic Press.

Wold, H. O. A. (1982). Soft Modeling: The Basic Design and Some Extensions. In K. G. Jöreskog & H. O. A. Wold (Eds.), *Systems Under Indirect Observations: Part II* (pp. 1–54). Amsterdam: North-Holland.

Wold, H. O. A. (1985). Partial Least Squares. In S. Kotz & N. L. Johnson (Eds.), *Encyclopedia of Statistical Sciences* (Vol. 6, pp. 581–591). New York: Wiley.

Yule, G. U. (1921). Critical Notice of "The Essentials of Mental Measurement" by W. Brown and G. H. Thomson. *British Journal of Psychology, 12*(1), 100–107.