Using PLS path modeling in new technology research: updated guidelines

Jörg Henseler

CTW, Laboratory of Design, Production & Management,
University of Twente, Enschede, The Netherlands

Geoffrey Hubona

Georgia R School, Chattanooga, Tennessee, USA, and

Pauline Ash Ray

Business Division, Thomas University, Thomasville, Georgia, USA

Abstract

Purpose – Partial least squares (PLS) path modeling is a variance-based structural equation modeling (SEM) technique that is widely applied in business and social sciences. Its ability to model composites and factors makes it a formidable statistical tool for new technology research. Recent reviews, discussions, and developments have led to substantial changes in the understanding and use of PLS. The paper aims to discuss these issues.

Design/methodology/approach – This paper aggregates new insights and offers a fresh look at PLS path modeling. It presents new developments, such as consistent PLS, confirmatory composite analysis, and the heterotrait-monotrait ratio of correlations.

Findings – PLS path modeling is the method of choice if a SEM contains both factors and composites. Novel tests of exact fit make a confirmatory use of PLS path modeling possible.

Originality/value – This paper provides updated guidelines of how to use PLS and how to report and interpret its results.

Keywords PLS, Structural equation modeling, Guidelines, Model test

Paper type Research paper

Introduction

Structural equation modeling (SEM) is a family of statistical techniques that has become very popular in business and social sciences. Its ability to model latent variables, to take into account various forms of measurement error, and to test entire theories makes it useful for a plethora of research questions.

Two types of SEM can be distinguished: covariance- and variance-based SEM. Covariance-based SEM estimates model parameters using the empirical variance-covariance matrix, and it is the method of choice if the hypothesized model consists of one or more common factors. In contrast, variance-based SEM first creates proxies as

linear combinations of observed variables, and then estimates the model parameters using these proxies. Variance-based SEM is the method of choice if the hypothesized model contains composites.

Among variance-based SEM methods, partial least squares (PLS) path modeling is regarded as the "most fully developed and general system" (McDonald, 1996, p. 240) and has been called a "silver bullet" (Hair *et al.*, 2011). PLS is widely used in information systems research (Marcoulides and Saunders, 2006), strategic management (Hair *et al.*, 2012a), marketing (Hair *et al.*, 2012b), and beyond. Its ability to model both factors and composites is appreciated by researchers across disciplines, and makes it a promising method particularly for new technology research and information systems research. Whereas factors can be used to model latent variables of behavioral research such as attitudes or personality traits, composites can be applied to model strong concepts (Höök and Löwgren, 2012), i.e. the abstraction of artifacts such as management instruments, innovations, or information systems. Consequently, PLS path modeling is a preferred statistical tool for success factor studies (Albers, 2010).

Not only has PLS and its use been subject of various reviews (cf. Hair *et al.*, 2012a, b), but just recently it has also undergone a series of serious examinations, and has been the target of heated scientific debates. Scholars have discussed the conceptual underpinnings (Rigdon, 2012, 2014; Sarstedt *et al.*, 2014) as well as the strengths and weaknesses (Rönkkö and Evermann, 2013; Henseler *et al.*, 2014; Aguirre-Urreta and Marakas, 2013; Rigdon *et al.*, 2014). As a fruitful outcome of these debates, substantial contributions to PLS emerged, such as bootstrap-based tests of overall model fit (Dijkstra and Henseler, 2015a), consistent PLS (PLSc) to estimate factor models (see Dijkstra and Henseler, 2015b), and the heterotrait-monotrait ratio of correlations (HTMT) as a new criterion for discriminant validity (see Henseler *et al.*, 2015). All these changes render the extant guidelines on PLS path modeling outdated, if not even invalid. Consequently, Rigdon (2014) recommends breaking the chains and forging ahead, which implies an urgent need for updated guidelines on why, when, and how to use PLS.

The purpose of our paper is manifold. First, it provides an updated view on what PLS actually is and which algorithmic steps it includes since the invention of PLSc. Second, it explains how to specify PLS path models, taking into account the nature of the measurement models (composite vs factor), model identification, sign indeterminacy, special treatments for categorical variables, and determination of sample size. Third, it explains how to assess and report PLS results, including the novel bootstrap-based tests of model fit, the SRMR as an approximate measure of model fit, the new reliability coefficient ρ_A, and the HTMT. Fourth, it sketches several ways of how to extend PLS analyses. Finally, it contrasts the understanding of PLS as presented in this paper with the traditional view, and discusses avenues for future developments.

The nature of PLS path modeling

The core of PLS is a family of alternating least squares algorithms that emulate and extend principal component analysis as well as canonical correlation analysis. The method was invented by Herman Wold (cf. 1974, 1982) for the analysis of high-dimensional data in a low-structure environment and has undergone various extensions and modifications. In its most modern appearance (cf. Dijkstra and Henseler, 2015a, b), PLS path modeling can be understood as a full-fledged SEM method that can handle both factor models and composite models for construct measurement, estimate recursive and non-recursive structural models, and conduct tests of model fit.

PLS path models are formally defined by two sets of linear equations: the measurement model (also called outer model) and the structural model (also called inner model). The measurement model specifies the relations between a construct and its observed indicators (also called manifest variables), whereas the structural model specifies the relationships between the constructs. Figure 1 depicts an example of a PLS path model.

PLS path models can contain two different forms of construct measurement: factor models or composite models (see Rigdon, 2012, for a nice comparison of both types of measurement models). The factor model hypothesizes that the variance of a set of indicators can be perfectly explained by the existence of one unobserved variable (the common factor) and individual random error. It is the standard model of behavioral research. In Figure 1, the exogenous construct ξ and the endogenous construct η are modeled as factors. In contrast, composites are formed as linear combinations of their respective indicators. The composite model does not impose any restrictions on the covariances between indicators of the same construct, i.e. it relaxes the assumption that all the covariation between a block of indicators is explained by a common factor. The composites serve as proxies for the scientific concept under investigation (Ketterlinus et al., 1989; Rigdon, 2012; Maraun and Halpin, 2008; Tenenhaus, 2008)[1]. The fact that composite models are less restrictive than factor models makes it likely that they have a higher overall model fit (Landis et al., 2000).

The structural model consists of exogenous and endogenous constructs as well as the relationships between them. The values of exogenous constructs are assumed to be given from outside the model. Thus, exogenous variables are not explained by other constructs in the model, and there must not be any arrows in the structural model that point to exogenous constructs. In contrast, endogenous constructs are at least partially explained by other constructs in the model. Each endogenous construct must have at least one arrow of the structural model pointing to it. The relationships between the constructs are usually assumed to be linear. The size and significance of path relationships is typically the focus of the scientific endeavors pursued in empirical research.
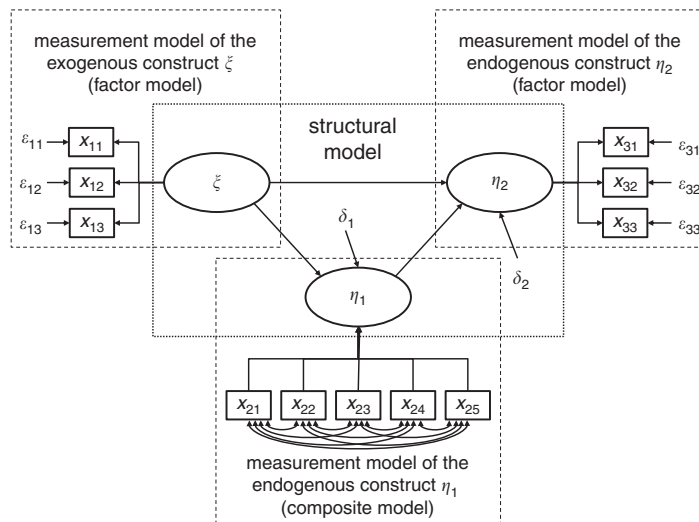

Figure 1.
PLS path model example

The estimation of PLS path model parameters happens in four steps: first, an iterative algorithm that determines composite scores for each construct; second, a correction for attenuation for those constructs that are modeled as factors (Dijkstra and Henseler, 2015b); third, parameter estimation; and finally, bootstrapping for inference testing.

Step 1: for each construct, the iterative PLS algorithm creates a proxy as a linear combination of the observed indicators. The indicator weights are determined such that each proxy shares as much variance as possible with the proxies of causally related constructs. The PLS algorithm can be viewed at as an approach to extend canonical correlation analysis to more than two sets of variables; it can emulate several of Kettenring's (1971) techniques for the canonical analysis of several sets of variables (Tenenhaus *et al.*, 2005). For a more detailed description of the algorithm see Henseler (2010). The major output of the first step are the proxies (i.e. composite scores), the proxy correlation matrix, and the indicator weights.

Step 2: correcting for attenuation is a necessary step if a model involves factors. As long as the indicators contain random measurement error, so will the proxies. Consequently, proxy correlations are typically underestimations of factor correlations. PLSc corrects for this tendency (Dijkstra and Henseler, 2015a, b) by dividing a proxy's correlations by the square root of its reliability (the so-called correction for attenuation). PLSc addresses the issue of what would the correlation between constructs be if there was no random measurement error? The major output of this second step is a consistent construct correlation matrix.

Step 3: once a consistent construct correlation matrix is available, it is possible to estimate the model parameters. If the structural model is recursive (i.e. there are no feedback loops), ordinary least squares (OLS) regression can be used to obtain consistent parameter estimates for the structural paths. In the case of non-recursive models, instrumental variable techniques such as two-stage least squares should be employed. Next to the path coefficient estimates, this third step can also provide estimates for loadings, indirect effects, total effects, and several model assessment criteria.

Step 4: finally, the bootstrap is applied in order to obtain inference statistics for all model parameters. The bootstrap is a non-parametric inferential technique which rests on the assumption that the sample distribution conveys information about the population distribution. Bootstrapping is the process of drawing a large number of re-samples with replacement from the original sample, and then estimating the model parameters for each bootstrap re-sample. The standard error of an estimate is inferred from the standard deviation of the bootstrap estimates.

The PLS path modeling algorithm has favorable convergence properties (Henseler, 2010). However, as soon as PLS path models involve common factors, there is the possibility of so-called Heywood cases (Krijnen *et al.*, 1998), meaning that one or more variances implied by the model would be negative. The occurrence of Heywood cases may be caused by an atypical or too-small sample, or the common factor structure may not hold for a particular set of indicators.

PLS path modeling is not as efficient as maximum likelihood covariance-based SEM. One possibility is to further minimize the discrepancy between the empirical and the model-implied correlation matrix, an approach followed by efficient PLS (see Bentler and Huang, 2014). Alternatively, one could embrace the notion that PLS is a limited-information estimator and is less affected by model misspecification in some subparts of a model (Antonakis *et al.*, 2010). Ultimately, there is no clear-cut resolution of the issues on this trade-off between efficiency and robustness with respect to model misspecification.

Model specification

The analysts must take care that the specified statistical model complies with the conceptual model intended to be tested, and further that the model complies with technical requirements such as identification, and with the data conforming to the required format and statistical power.

Typically, the structural model is theory based and is the prime focus of the research question and/or research hypotheses. The specification of the structural model addresses two questions: Which constructs should be included in the model? And how are they hypothesized to be interrelated? That is, what are the directions and strengths of the causal influences between and among the latent constructs? In general, analysts should keep in mind that the constructs specified in a model are only proxies, and that there will always be a validity gap between these proxies and the theoretical concepts that are the intended modeling target (Rigdon, 2012). The paths, specified as arrows in a PLS model, represent directional linear relationships between proxies. The structural model, and the indicated relationships among the latent constructs, is regarded as separate from the measurement model.

The specification of the measurement model entails decisions for composite or factor models and the assignment of indicators to constructs. Factor models are the predominant measurement model for behavioral constructs such as attitudes or personality traits. Factor models are strongly linked to true score theory (McDonald, 1999), the most important measurement paradigm in behavioral sciences. If a construct has this background and random measurement error is likely to be an issue, analysts should choose the factor model. Composites help model emergent constructs, for which elements are combined to form a new entity. Composites can be applied to model strong concepts (Höök and Löwgren, 2012), i.e. the abstraction of artifacts (man-made objects). Typical artifacts in new technology research would include innovations, technologies, systems, processes, strategies, management instruments, or portfolios. Whenever a model contains this type of construct it is preferable to opt to use a composite model.

Measurement models of PLS path models may appear less detailed than those of covariance-based SEM, but in fact some specifications are implicit and are not visualized. For instance, neither the unique indicator errors (nor their correlations) of factor models nor the correlations between indicators of composite models are drawn. Because PLS currently does not allow to either constrain these parameters nor to free the error correlations of factor models, by convention these model elements are not drawn. No matter which type of measurement is chosen to measure a construct, PLS requires that there is at least one indicator available. Constructs without indicators, so-called phantom variables (Rindskopf, 1984), cannot be included in PLS path models.

In some PLS path modeling software (e.g. SmartPLS and PLS-Graph), the depicted direction of arrows in the measurement model does not indicate whether a factor or composite model is estimated, but whether correlation weights (Mode A, represented by arrows pointing from a construct to its indicators) or regression weights (Mode B, represented by arrows pointing from indicators to their construct) shall be used to create the proxy. In both cases PLS will estimate a composite model. Indicator weights estimated by Mode B are consistent (Dijkstra, 2010) whereas indicators weights estimated by Mode A are not, but the latter excel in out-of-sample prediction (Rigdon, 2012).

Some model specifications are made automatically and cannot be manually changed: measurement errors are assumed to be uncorrelated with all other variables and errors in the model; structural disturbance terms are assumed to be orthogonal to their predictor

variables as well as to each other[2]; correlations between exogenous variables are free. Because these specifications hold across models, it has become customary not to draw them in PLS path models.

Identification has always been an important issue for SEM, although it has been neglected in the realm of PLS path modeling in the past. It refers to the necessity to specify a model such that only one set of estimates exists that yields the same model-implied correlation matrix. It is possible that a complete model is unidentified, but also only parts of a model can be unidentified. In general, it is not possible to derive useful conclusions from unidentified (parts of) models. In order to achieve identification, PLS fixes the variance of factors and composites to one. An important requirement of composite models is a so-called nomological net. It means that composites cannot be estimated in isolation, but need at least one other variable (either observed or latent) to have a relation with. Since PLS also estimates factor models via composites, this requirement extends to all factor models estimated using PLS. If a factor model has exactly two indicators, it does not matter which form of SEM is used – a nomological net is then required to achieve identification. If a construct is only measured by one indicator, one speaks of single-indicator measurement (Diamantopoulos *et al.*, 2012). The construct scores are then identical to the standardized indicator values. In this case it is not possible to determine the amount of random measurement error in this indicator. If an indicator is error-prone, the only possibility to account for the error is to utilize external knowledge about the reliability of this indicator to manually define the indicator's reliability.

A typical characteristic of SEM and factor-analytical tools in general is sign indeterminacy, in which the weight or loading estimates for a factor or a composite can only be determined jointly for their value but not for their sign. For example, if a factor is extracted from the strongly negatively correlated customer satisfaction indicators "How satisfied are you with provider X?" and "How much does provider X differ from an ideal provider?" The method cannot "know" whether the extracted factor should correlate positively with the first or with the second indicator. Depending on the sign of the loadings, the meaning of the factor would either be "customer satisfaction" or "customer non-satisfaction." To avoid this ambiguity, it has become practice in SEM to determine one particular indicator per construct with which the construct scores are forced to correlate positively[3]. Since this indicator dictates the orientation of the construct, it is called the "dominant indicator." While in covariance-based SEM this dominant indicator also dictates the construct's variance, in PLS path modeling the construct variance is simply set to one.

Like multiple regression, PLS path modeling requires metric data for the dependent variables. Dependent variables are the indicators of the factor model(s) as well as the endogenous constructs. Quasi-metric data stemming from multi-point scales such as Likert scales or semantic differential scales is also acceptable as long as the scale points can be assumed to be equidistant. To some extent it is also possible to include categorical variables in a model. Categorical variables are particularly relevant for analyzing experiments (cf. Streukens *et al.*, 2010) or for control variables such as industry (Braojos-Gomez *et al.*, 2015) or ownership structure (Chen *et al.*, 2015). Figure 2 illustrates how a categorical variable "marital status" would be included in a PLS path model. If a categorical variable has only two levels (i.e. it is dichotomous), it can serve immediately as a construct indicator. If a categorical variable has more than two levels, it should be transformed into as many dummy variables as there are levels. A composite model is formed out of all but one dummy variable. The remaining

Figure 2.
Including a
categorical control
variable in a PLS
path model
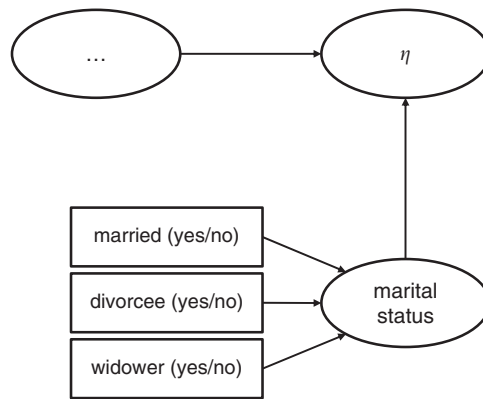
Note: Marital status with the four categories
"unmarried", "married", "divorcee", "widower";
the reference category is "unmarried"

dummy variable characterizes the reference level. Preferably, categorical variables should only play the role of exogenous variables in a structural model.

Sample size plays a dual role, namely, technically and in terms of inference statistics. Technically, the number of observations must be high enough that the regressions that form part of the PLS algorithm do not evoke singularities. It can thus be that the number of parameters or the number of variables in a model exceeds the number of observations. Inference statistics become relevant if an analyst wants to generalize from a sample to a population. The larger the sample size, the smaller the confidence intervals of the model's parameter estimates, and the smaller the chance that a parameter estimate's deviation from zero is due to sampling variation. Moreover, a larger sample size increases the likelihood to detect model misspecification (see fourth section for PLS' tests of model fit). Hence, a larger sample size increases the rigor to falsify the model in the Popperian sense, but at the same time the likelihood increases that a model gets rejected due to minor and hardly relevant aspects. The statistical power of PLS should not be expected to supersede that of covariance-based SEM[4]. Consequently, there is no reason to prefer PLS over other forms of SEM with regard to inference statistics. In research practice, there are typically many issues that have an impact on the final sample size. One important consideration should be the statistical power, i.e. the likelihood to find an effect in the sample if it indeed exists in the population. Optimally, researchers make use of Monte Carlo simulations to quantify the statistical power achieved at a certain sample size (for a tutorial, see Aguirre-Urreta and Rönkkö, 2015).

Assessing and reporting PLS analyses

PLS path modeling can be used both for explanatory and predictive research. Depending on the analyst's aim – either explanation or prediction – the model assessment will be different. If the analyst's aim is to predict, the assessment should focus on blindfolding (Tenenhaus *et al.*, 2005) and the model's performance with regard to holdout samples. However, since prediction-orientation still tends to be scarce in business research (Shmueli and Koppius, 2013), in the remainder we will focus on model assessment if the analyst's aim is explanation.

PLS path modeling results can be assessed globally (i.e. for the overall model) and locally (for the measurement models and the structural model). For a long time it was said that PLS path modeling does not optimize any global scalar and therefore does not allow for global model assessment. However, because PLS in the form as described above provides consistent estimates for factor and composite models, it is possible to meaningfully compare the model-implied correlation matrix with the empirical correlation matrix, which opens up the possibility for the assessment of global model fit.

The overall goodness-of-fit (GoF) of the model should be the starting point of model assessment. If the model does not fit the data, the data contains more information than the model conveys. The obtained estimates may be meaningless, and the conclusions drawn from them become questionable. The global model fit can be assessed in two non-exclusive ways: by means of inference statistics, i.e. so-called tests of model fit, or through the use of fit indices, i.e. an assessment of approximate model fit. In order to have some frame of reference, it has become customary to determine the model fit both for the estimated model and for the saturated model. Saturation refers to the structural model, which means that in the saturated model all constructs correlate freely.

PLS path modeling's tests of model fit rely on the bootstrap to determine the likelihood of obtaining a discrepancy between the empirical and the model-implied correlation matrix that is as high as the one obtained for the sample at hand if the hypothesized model was indeed correct (Dijkstra and Henseler, 2015a). Bootstrap samples are drawn from modified sample data. This modification entails an orthogonalization of all variables and a subsequent imposition of the model-implied correlation matrix. In covariance-based SEM, this approach is known as Bollen-Stine bootstrap (Bollen and Stine, 1992). If more than 5 percent (or a different percentage if an α-level different from 0.05 is chosen) of the bootstrap samples yield discrepancy values above the ones of the actual model, it is not that unlikely that the sample data stems from a population that functions according to the hypothesized model. The model thus cannot be rejected. There is more than one way to quantify the discrepancy between two matrices, for instance the maximum likelihood discrepancy, the geodesic discrepancy d_G, or the unweighted least squares discrepancy d_{ULS} (Dijkstra and Henseler, 2015a), and so there are several tests of model fit. Monte Carlo simulations confirm that the tests of model fit can indeed discriminate between well-fitting and ill-fitting models (Henseler *et al.*, 2014). More precisely, both measurement model misspecification and structural model misspecification can be detected through the tests of model fit (Dijkstra and Henseler, 2014). Because it is possible that different tests have different results, a transparent reporting practice would always include several tests.

Next to conducting the tests of model fit it is also possible to determine the approximate model fit. Approximate model fit criteria help answer the question how substantial the discrepancy between the model-implied and the empirical correlation matrix is. This question is particularly relevant if this discrepancy is significant. Currently, the only approximate model fit criterion implemented for PLS path modeling is the standardized root mean square residual (SRMR) (Hu and Bentler, 1998, 1999). As can be derived from its name, the SRMR is the square root of the sum of the squared differences between the model-implied and the empirical correlation matrix, i.e. the Euclidean distance between the two matrices. A value of 0 for SRMR would indicate a perfect fit and generally, an SRMR value less than 0.05 indicates an acceptable fit (Byrne, 2008). A recent simulation study shows that even entirely correctly specified model can yield SRMR values of 0.06 and higher (Henseler *et al.*, 2014). Therefore, a cut-off value of 0.08 as proposed by Hu and Bentler (1999) appears to be more adequate

for PLS path models. Another useful approximate model fit criterion could be the Bentler-Bonett index or normed fit index (NFI) (Bentler and Bonett, 1980). The suggestion to use the NFI in connection with PLS path modeling can be attributed to Lohmöller (1989). For factor models, NFI values above 0.90 are considered as acceptable (Byrne, 2008). For composite models, thresholds for the NFI are still to be determined. Because the NFI does not penalize for adding parameters, it should be used with caution for model comparisons. In general, the usage of the NFI is still rare[5]. Another promising approximate model fit criterion is the root mean square error correlation (RMS$_{theta}$) (see Lohmöller, 1989). A recent simulation study (Henseler et $al.$, 2014) provides evidence that the RMS$_{theta}$ can indeed distinguish well-specified from ill-specified models. However, thresholds for the RMS$_{theta}$ are yet to be determined, and PLS software still needs to implement this approximate model fit criterion. Note that early suggestions for PLS-based GoF measures such as the "goodness-of-fit" (see Tenenhaus et $al.$, 2004) or the "relative goodness-of-fit" (proposed by Esposito Vinzi et $al.$, 2010) are – in opposite to what their name might suggest – not informative about the goodness of model fit (Henseler and Sarstedt, 2013; Henseler et $al.$, 2014). Consequently, there is no reason to evaluate and report them if the analyst's aim is to test or to compare models.

If the specified measurement (or outer) model does not possess minimum required properties of acceptable reliability and validity, then the structural (inner) model estimates become meaningless. That is, a necessary condition to even proceed to assess the "goodness" of the inner structural model is that the outer measurement model has already demonstrated acceptable levels of reliability and validity. There must be a sound measurement model before one can begin to assess the "goodness" of the inner structural model or to rely on the magnitude, direction, and/or statistical strength of the structural model's estimated parameters. Factor and composite models are assessed in a different way.

Factor models can be assessed in various ways. The bootstrap-based tests of overall model fit can indicate whether the data are coherent with a factor model, i.e. it represents a confirmatory factor analysis. In essence, the test of model fit provides an answer to the question "Does empirical evidence speak against that the factor exists?" This quest for truth illustrates that testing factor model is rooted in the positivist research paradigm. Once the test of overall model fit has not provided evidence against the existence of a factor[6], several questions with regard to the factor structure emerge: does the data support a factor structure at all? Can a factor unanimously be extracted? How well has this factor been measured? Note that tests of overall model fit cannot answer these questions; in particular, entirely uncorrelated empirical variables do not necessarily lead to the rejection of the factor model. To answer these questions one should rather rely on several local assessment criteria with regard to the reliability and validity of measurement.

The amount of random error in construct scores should be acceptable, or in other words: the reliability of construct scores should be sufficiently high. Nunnally and Bernstein (1994) recommend a minimum reliability of 0.7. The most important reliability measure for PLS is ρ_A (Dijkstra and Henseler, 2015b); it currently is the only consistent reliability measure for PLS construct scores. Most PLS software also provides a measure of composite reliability (also called Dillon-Goldstein's ρ, factor reliability, Jöreskog's ρ, ω, or ρ_c) as well as Cronbach's α. Both refer to sum scores, not construct scores. In particular, Cronbach's α typically underestimates the true reliability, and should therefore only be regarded as a lower boundary to the reliability (Sijtsma, 2009).

The measurement of factors should also be free from systematic measurement error. This quest for validity can be fulfilled in several non-exclusive ways. First, a factor should be unidimensional, a characteristic examined through convergent validity. The dominant measure of convergent validity is the average variance extracted (AVE) (Fornell and Larcker, 1981)[7]. If the first factor extracted from a set of indicators explains more than one half of their variance, there cannot be any second, equally important factor. An AVE of 0.5 or higher is therefore regarded as acceptable. A somewhat more liberal criterion was proposed by Sahmer et al. (2006): they find evidence for unidimensionality as long as a factor explains significantly more variance than the second factor extracted from the same indicators. Second, each pair of factors that stand in for theoretically different concepts should also statistically be different, which raises the question of discriminant validity. Two criteria have been shown to be informative about discriminant validity (Voorhees et al., forthcoming): the Fornell-Larcker criterion (proposed by Fornell and Larcker, 1981) and the HTMT (developed by Henseler et al., 2015). The Fornell-Larcker criterion says that a factor's AVE should be higher than its squared correlations with all other factors in the model. The HTMT is an estimate for the factor correlation (more precisely, an upper boundary). In order to clearly discriminate between two factors, the HTMT should be significantly smaller than one. Third, the cross-loadings should be assessed to make sure that no indicator is incorrectly assigned to a wrong factor.

The assessment of composite models is somewhat less developed. Again, the major point of departure should be the tests of model fit. The tests of model fit for the saturated model provide evidence for the external validity of the composites. Henseler et al. (2014) call this step a "confirmatory composite analysis." For composite models, the major research question is "Does it make sense to create this composite?" This different question shows that testing composite models follows a different research paradigm, namely, pragmatism (Henseler, 2015). Once confirmatory composite analysis has provided support for the composite, it can be analyzed further. One follow-up suggests itself: How is the composite made? Do all the ingredients contribute significantly and substantially? To answer these questions, an analyst should assess the sign and the magnitude of the indicator weights as well as their significance. Particularly if indicators weights have unexpected signs or are insignificant, this can be due to multicollinearity. It is therefore recommendable to assess the variance inflation factor (VIF) of the indicators. VIF values much higher than one indicate that multicollinearity might play a role.

Once the measurement model is deemed to be of sufficient quality, the analyst can proceed and assess the structural model. If OLS is used for the structural model, the endogenous constructs' R^2 values would be the point of departure. They indicate the percentage of variability accounted for by the precursor constructs in the model. The adjusted R^2 values take into account model complexity and sample size, and are thus helpful to compare different models or the explanatory power of a model across different data sets.

If the analyst's aim is to generalize from a sample to a population, the path coefficients should be evaluated for significance. Inference statistics include the empirical bootstrap confidence intervals as well as one-sided or two-sided p-values. We recommend to use 4,999 bootstrap samples. This number is sufficiently close to infinity for usual situations, is tractable with regard to computation time, and allows for an unanimous determination of empirical bootstrap confidence intervals (for instance, the 2.5 percent (97.5 percent) quantile would be the 125th (4,875th) element of the sorted list of bootstrap values). A path coefficient is regarded as significant (i.e. unlikely to purely result from sampling error) if its confidence interval does not include the value of zero or if the p-value is below

the pre-defined α-level. Despite strong pleas for the use of confidence intervals (Cohen, 1994), reporting p-values still seems to be more common in business research.

For the significant effects it makes sense to quantify how substantial they are, which can be accomplished by assessing their effect size f^2. f^2 values above 0.35, 0.15, and 0.02 can be regarded as strong, moderate, and weak, respectively (Cohen, 1988). The path coefficients are essentially standardized regression coefficients, which can be assessed with regard to their sign and their absolute size. They should be interpreted as the change in the dependent variable if the independent variable is increased by one and all other independent variables remain constant. Indirect effects and their inference statistics are important for mediation analysis (Zhao *et al.*, 2010), and total effects are useful for success factor analysis (Albers, 2010). Table I sums up the discussed criteria for model assessment.

Extensions

PLS path modeling as described so far analyzes linear relationships between factors or composites of observed indicator variables. There are many ways how this rather basic model can be extended.

Assessment	Criterion
Overall model	
Test of model fit (estimated model)	SRMR < 95% bootstrap quantile (HI95 of SRMR)
	d_{ULS} < 95% bootstrap quantile (HI95 of d_{ULS})
	d_G < 95% bootstrap quantile (HI95 of d_G)
Approximate model fit (estimated model)	SRMR < 0.08
Measurement model	
Confirmatory composite and/or factor analysis (saturated model)	SRMR < 95% bootstrap quantile (HI95 of SRMR)
	d_{ULS} < 95% bootstrap quantile (HI95 of d_{ULS})
	d_G < 95% bootstrap quantile (HI95 of d_G)
Approximate model fit (saturated model)	SRMR < 0.08
Internal consistency reliability	Dijkstra-Henseler's ρ_A > 0.7
	Dillon-Goldstein's ρ_c > 0.7
	Cronbach's α > 0.7
Convergent validity	AVE > 0.5
Discriminant validity	HTMT significantly smaller than 1
	Fornell-Larcker criterion
	Loadings exceed cross-loadings
Structural model	
Endogenous variables	R^2, adjusted R^2
Direct effects	Path coefficient (absolute size, sign)
	Significance (p-value, confidence interval)
	Effect size
Indirect effects	Coefficient (absolute size, sign)
	Significance (p-value, confidence interval)
Total effects	Coefficient (absolute size, sign)
	Significance (p-value, confidence interval)

Table I.
Assessment of PLS path modeling results in explanatory research settings

A first extension is to depart from the assumption of linearity. Researchers have developed approaches to include non-linear relationships into the structural model. In particular, interaction effects and quadratic effects can be easily analyzed by means of some rudimentary extensions to the standard PLS path modeling setup (Dijkstra and Henseler, 2011; Henseler and Fassott, 2010; Henseler et al., 2012; Henseler and Chin, 2010; Dijkstra and Schermelleh-Engel, 2014). Interaction effects pay tribute to the fact that not all individuals function according to the same mechanism, but that the strength of relationships depends on contingencies.

Next to interaction effects, there are more comprehensive tools to take into account the heterogeneity between individuals. Heterogeneity can be observed, i.e. it can be traced back to an identified variable, or unobserved, i.e. there is no a priori explanation for why an individual's mechanism would differ from others. Because incorrectly assuming that all individuals function according to the same mechanism represents a validity thread (Becker et al., 2013b), several PLS-based approaches to discover unobserved heterogeneity have been proposed. Prominent examples include finite mixture PLS (Ringle et al., 2010a, c), PLS prediction-oriented segmentation (Becker et al., 2013b), and PLS genetic algorithm segmentation (Ringle et al., 2010b, 2014). In order to assess observed heterogeneity, analysts should make use of multigroup analysis (Sarstedt et al., 2011). No matter whether heterogeneity is observed or unobserved, another concern for the analysts must be not to confound heterogeneity in the structural model with variation in measurement. Particularly in cross-cultural research is has therefore become a common practice to assess the measurement model invariance before drawing conclusions about structural model heterogeneity. There is a plethora of papers discussing how to assess the measurement invariance of factor models (see e.g. French and Finch, 2006), there is only one approach for assessing the measurement invariance of composite models (Henseler et al., forthcoming).

Discussion
The plethora of discussions and developments around PLS path modeling called for a fresh look at this technique as well as new guidelines. As important aspect of this endeavor, we provide an answer the question "What has changed?" This answer is given in Table II, which contrasts traditional and modern perspectives on PLS. It is particularly helpful for researchers who have been educated in PLS path modeling in the past, and who would like to update their understanding of the method.

The fact that PLS today strongly differs from how it used to be has also implications for the users of PLS software. They should verify that they use very current versions of PLS software such as SmartPLS, which have implemented the newest developments in the PLS field. Alternatively, they may want to use ADANCO (Henseler and Dijkstra, 2015), a new software for variance-based SEM, which also includes PLS path modeling.

The modularity of PLS path modeling as introduced in the second section opens up the possibility of replacing one or more steps by other approaches. For instance, the least squares estimators of the third step could be replaced by neural networks (Buckler and Hennig-Thurau, 2008; Turkyilmaz et al., 2013). One could even replace the PLS algorithm in Step 1 by alternative indicator weight generators, such as principal component analysis (Tenenhaus, 2008), generalized structured component analysis (Hwang and Takane, 2004; Henseler, 2012), regularized generalized canonical correlation analysis (Tenenhaus and Tenenhaus, 2011), or even plain sum scores. Because in these instances the iterative PLS algorithm would not serve as eponym, one could not speak of PLS path modeling any more. However, it still would be variance-based SEM.

Traditional view on PLS	Modern view on PLS
PLS has some but not all abilities of structural equation modeling	PLS is a full-fledged structural equation modeling approach
PLS can estimate formative (using Mode B) and reflective measurement models (using Mode A)	PLS can consistently estimate composite models (using Mode B) and factor models (using consistent PLS for the latter)
Identification is not an issue for PLS	To ensure identification, analysts must provide a nomological net for each multi-item construct
PLS path models must be recursive	PLS path models can contain feedback loops or take into account endogeneity if an adequate estimator is used for the structural model. A sufficient number of exogenous variables must be available
PLS needs fewer observations than other SEM techniques	PLS does not need fewer observations than other techniques when it comes to inference statistics. Analysts should ensure sufficient statistical power and representativeness of data
In contrast to other SEM techniques, PLS does not rely on the assumption of normality	With regard to assumptions made for the estimation of parameters, PLS does not differ from other SEM techniques. For inference statistics, PLS applies a non-parametric technique, namely, bootstrapping, which can equally be applied by other SEM techniques
PLS only permits local model assessment by means of certain criteria	PLS path models can and should be assessed globally by means of tests of model fit and approximate measures of model fit. Additionally, models should be locally assessed
The reliability of PLS construct scores is indicated by Cronbach's α and/or composite reliability	The reliability coefficient ρ_A is a consistent estimate of the reliability of PLS construct scores; composite reliability (based on consistent loadings) is a consistent estimate of the reliability of sum scores
Discriminant validity should be assessed by comparing each construct's average variance extracted with its squared construct correlations	Discriminant validity should be assessed by means of the heterotrait-monotrait ratio of correlations (HTMT) and by comparing each construct's average variance extracted (based on consistent loadings) with its squared consistent construct correlations
Bootstrapping should be conducted in combination with sign change correction in order to avoid inflated standard errors	For each construct, a dominant indicator should be defined in order to avoid sign indeterminacy

Table II.
Contrasting traditional and modern perspectives on PLS

Finally, recent research confirms that PLS serves as a promising technique for prediction purposes (Becker *et al.*, 2013a). Both measurement models and structural models can be assessed with regard to their predictive validity. Blindfolding is the standard approach used to examine if the model or a single effect of it can predict values of reflective indicators. It is already widely applied (Hair *et al.*, 2012b; Ringle *et al.*, 2012). Criteria for the predictive capability of structural models have been proposed (cf. Chin, 2010), but still need to disseminate. We anticipate that once business and social science researchers' interest in prediction becomes more pronounced, PLS will face an additional substantial increase in popularity.

Notes

1. Note that also factors are nothing else than proxies (Rigdon, 2012).

2. This assumption should be relaxed in case of non-recursive models (Dijkstra and Henseler, 2015a).

3. Automated sign-change procedures such as "individual sign change" or "construct level sign change" should be regarded as deprecated.

4. An allegedly higher statistical power of PLS (Reinartz *et al.*, 2009) can be traced back to model misspecification, namely, making use of a composite model although the factor model would have been true (Goodhue *et al.*, 2011).

5. For an application of the NFI, see Ziggers and Henseler (forthcoming).

6. Interestingly, the methodological literature on factor models is quite silent about what to do if the test speaks against a factor model. Some researchers suggest to consider the alternative of a composite model, because it is less restrictive (Henseler *et al.*, 2014) and not subject to factor indeterminacy (Rigdon, 2012).

7. The AVE must be calculated based on consistent loadings, otherwise the assessment of convergent and discriminant validity based on the AVE is meaningless.

References

Aguirre-Urreta, M. and Rönkkö, M. (2015), "Sample size determination and statistical power analysis in PLS using R: an annotated tutorial", *Communications of the Association for Information Systems*, Vol. 36 No. 3, pp. 33-51.

Aguirre-Urreta, M.I. and Marakas, G.M. (2013), "Research note – partial least squares and models with formatively specified endogenous constructs: a cautionary note", *Information Systems Research*, Vol. 25 No. 4, pp. 761-778.

Albers, S. (2010), "PLS and success factor studies in marketing", in Esposito Vinzi, V., Chin, W.W., Henseler, J. and Wang, H. (Eds), *Handbook of Partial Least Squares*, Springer, Berlin, pp. 409-425.

Antonakis, J., Bendahan, S., Jacquart, P. and Lalive, R. (2010), "On making causal claims: a review and recommendations", *The Leadership Quarterly*, Vol. 21 No. 6, pp. 1086-1120.

Becker, J.-M., Rai, A. and Rigdon, E.E. (2013a), "Predictive validity and formative measurement in structural equation modeling: embracing practical relevance", *International Conference on Information Systems*, Milan, December 15-18.

Becker, J.-M., Rai, A., Ringle, C.M. and Völckner, F. (2013b), "Discovering unobserved heterogeneity in structural equation models to avert validity threats", *MIS Quarterly*, Vol. 37 No. 3, pp. 665-694.

Bentler, P.M. and Bonett, D.G. (1980), "Significance tests and goodness of fit in the analysis of covariance structures", *Psychological Bulletin*, Vol. 88 No. 3, pp. 588-606.

Bentler, P.M. and Huang, W. (2014), "On components, latent variables, PLS and simple methods: reactions to Rigdon's rethinking of PLS", *Long Range Planning*, Vol. 47 No. 3, pp. 138-145.

Bollen, K.A. and Stine, R.A. (1992), "Bootstrapping goodness-of-fit measures in structural equation models", *Sociological Methods & Research*, Vol. 21 No. 2, pp. 205-229.

Braojos-Gomez, J., Benitez-Amado, J. and Llorens-Montes, F.J. (2015), "How do small firms learn to develop a social media competence?", *International Journal of Information Management*, Vol. 35 No. 4, pp. 443-458.

Buckler, F. and Hennig-Thurau, T. (2008), "Identifying hidden structures in marketing's structural models through universal structure modeling: an explorative Bayesian neural network complement to LISREL and PLS", *Marketing-Journal of Research and Management*, Vol. 4 No. 2, pp. 47-66.

Byrne, B.M. (2008), *Structural Equation Modeling with EQS: Basic Concepts, Applications, and Programming*, Psychology Press, New York, NY.

Chen, Y., Wang, Y., Nevo, S., Benitez-Amado, J. and Kou, G. (2015), "IT capabilities and product innovation performance: the roles of corporate entrepreneurship and competitive intensity", *Information & Management*, Vol. 52 No. 6, pp. 643-657.

Chin, W.W. (2010), "Bootstrap cross-validation indices for PLS path model assessment", in Esposito Vinzi, V., Chin, W.W., Henseler, J. and Wang, H. (Eds), *Handbook of Partial Least Squares: Concepts, Methods and Applications*, Springer, Heidelberg, Dordrecht, London and New York, NY, pp. 83-97.

Cohen, J. (1988), *Statistical Power Analysis for the Behavioral Sciences*, Lawrence Erlbaum, Mahwah, NJ.

Cohen, J. (1994), "The earth is round (p < 0.05)", *American Psychologist*, Vol. 49 No. 12, pp. 997-1003.

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P. and 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*, Vol. 40 No. 3, pp. 434-449.

Dijkstra, T.K. (2010), "Latent variables and indices: Herman Wold's basic design and partial least squares", in Esposito Vinzi, V., Chin, W.W., Henseler, J. and Wang, H. (Eds), *Handbook of Partial Least Squares: Concepts, Methods and Applications*, Springer, Heidelberg, Dordrecht, London and New York, NY, pp. 23-46.

Dijkstra, T.K. and Henseler, J. (2011), "Linear indices in nonlinear structural equation models: best fitting proper indices and other composites", *Quality & Quantity*, Vol. 45 No. 6, pp. 1505-1518.

Dijkstra, T.K. and Henseler, J. (2014), "Assessing and testing the goodness-of-fit of PLS path models", *3rd VOC Conference, Leiden, May 9*.

Dijkstra, T.K. and Henseler, J. (2015a), "Consistent and asymptotically normal PLS estimators for linear structural equations", *Computational Statistics & Data Analysis*, Vol. 81 No. 1, pp. 10-23.

Dijkstra, T.K. and Henseler, J. (2015b), "Consistent partial least squares path modeling", *MIS Quarterly*, Vol. 39 No. 2, pp. 297-316.

Dijkstra, T.K. and Schermelleh-Engel, K. (2014), "Consistent partial least squares for nonlinear structural equation models", *Psychometrika*, Vol. 79 No. 4, pp. 585-604.

Esposito Vinzi, V., Trinchera, L. and Amato, S. (2010), "PLS path modeling: from foundations to recent developments and open issues for model assessment and improvement", in Esposito Vinzi, V., Chin, W.W., Henseler, J. and Wang, H. (Eds), *Handbook of Partial Least Squares: Concepts, Methods and Applications*, Springer, Berlin, pp. 47-82.

Fornell, C. and Larcker, D.F. (1981), "Evaluating structural equation models with unobservable variables and measurement error", *Journal of Marketing Research*, Vol. 18 No. 1, pp. 39-50.

French, B.F. and Finch, W.H. (2006), "Confirmatory factor analytic procedures for the determination of measurement invariance", *Structural Equation Modeling*, Vol. 13 No. 3, pp. 378-402.

Goodhue, D.L., Lewis, W. and Thompson, R.L. (2011), "A dangerous blind spot in IS research: false positives due to multicollinearity combined with measurement error", *AMCIS 2011, Detroit, MI, August 4-8*.

Hair, J.F., Ringle, C.M. and Sarstedt, M. (2011), "PLS-SEM: indeed a silver bullet", *Journal of Marketing Theory and Practice*, Vol. 18 No. 2, pp. 139-152.

Hair, J.F., Sarstedt, M., Pieper, T.M. and 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*, Vol. 45 Nos 5/6, pp. 320-340.

Hair, J.F., Sarstedt, M., Ringle, C.M. and 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*, Vol. 40 No. 3, pp. 414-433.

Henseler, J. (2010), "On the convergence of the partial least squares path modeling algorithm", *Computational Statistics*, Vol. 25 No. 1, pp. 107-120.

Henseler, J. (2012), "Why generalized structured component analysis is not universally preferable to structural equation modeling", *Journal of the Academy of Marketing Science*, Vol. 40 No. 3, pp. 402-413.

Henseler, J. (2015), "Is the whole more than the sum of its parts? On the interplay of marketing and design research", Inaugural lecture held on 30 April 2015, University of Twente, Enschede.

Henseler, J. and Chin, W.W. (2010), "A comparison of approaches for the analysis of interaction effects between latent variables using partial least squares path modeling", *Structural Equation Modeling*, Vol. 17 No. 1, pp. 82-109.

Henseler, J. and Dijkstra, T.K. (2015), "ADANCO 2.0", Composite Modeling, Kleve, available at: www.compositemodeling.com (accessed December 14, 2015).

Henseler, J. and Fassott, G. (2010), "Testing moderating effects in PLS path models: an illustration of available procedures", in Esposito Vinzi, V., Chin, W.W., Henseler, J. and Wang, H. (Eds), *Handbook of Partial Least Squares: Concepts, Methods and Applications*, Springer, Berlin, pp. 713-735.

Henseler, J. and Sarstedt, M. (2013), "Goodness-of-fit indices for partial least squares path modeling", *Computational Statistics*, Vol. 28 No. 2, pp. 565-580.

Henseler, J., Ringle, C.M. and Sarstedt, M. (2015), "A new criterion for assessing discriminant validity in variance-based structural equation modeling", *Journal of the Academy of Marketing Science*, Vol. 43 No. 1, pp. 115-135.

Henseler, J., Ringle, C.M. and Sarstedt, M. (forthcoming), "Testing measurement invariance of composites using partial least squares", *International Marketing Review* (in print).

Henseler, J., Fassott, G., Dijkstra, T.K. and Wilson, B. (2012), "Analysing quadratic effects of formative constructs by means of variance-based structural equation modelling", *European Journal of Information Systems*, Vol. 21 No. 1, pp. 99-112.

Henseler, J., Dijkstra, T.K., Sarstedt, M., Ringle, C.M., Diamantopoulos, A., Straub, D.W., Ketchen, D.J. Jr, Hair, J.F., Hult, G.T.M. and Calantone, R.J. (2014), "Common beliefs and reality about PLS: comments on Rönkkö & Evermann (2013)", *Organizational Research Methods*, Vol. 17 No. 2, pp. 182-209.

Höök, K. and Löwgren, J. (2012), "Strong concepts: intermediate-level knowledge in interaction design research", *ACM Transactions on Computer-Human Interaction (TOCHI)*, Vol. 19 No. 3, Article 23.

Hu, L. and Bentler, P.M. (1998), "Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification", *Psychological Methods*, Vol. 3 No. 4, pp. 424-453.

Hu, L.-T. and Bentler, P.M. (1999), "Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives", *Structural Equation Modeling*, Vol. 6 No. 1, pp. 1-55.

Hwang, H. and Takane, Y. (2004), "Generalized structured component analysis", *Psychometrika*, Vol. 69 No. 1, pp. 81-99.

Kettenring, J.R. (1971), "Canonical analysis of several sets of variables", *Biometrika*, Vol. 58 No. 3, pp. 433-451.

Ketterlinus, R.D., Bookstein, F.L., Sampson, P.D. and Lamb, M.E. (1989), "Partial least squares analysis in developmental psychopathology", *Development and Psychopathology*, Vol. 1 No. 2, pp. 351-371.

Krijnen, W.P., Dijkstra, T.K. and Gill, R.D. (1998), "Conditions for factor (in)determinacy in factor analysis", *Psychometrika*, Vol. 63 No. 4, pp. 359-367.

Landis, R.S., Beal, D.J. and Tesluk, P.E. (2000), "A comparison of approaches to forming composite measures in structural equation models", *Organizational Research Methods*, Vol. 3 No. 2, pp. 186-207.

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

McDonald, R.P. (1996), "Path analysis with composite variables", *Multivariate Behavioral Research*, Vol. 31 No. 2, pp. 239-270.

McDonald, R.P. (1999), *Test Theory: A Unified Treatment*, Lawrence Erlbaum, Mahwah, NJ.

Maraun, M.D. and Halpin, P.F. (2008), "Manifest and latent variates", *Measurement: Interdisciplinary Research and Perspectives*, Vol. 6 Nos 1/2, pp. 113-117.

Marcoulides, G.A. and Saunders, C. (2006), "PLS: a silver bullet?", *MIS Quarterly*, Vol. 30 No. 2, pp. iii-ix.

Nunnally, J.C. and Bernstein, I.H. (1994), *Psychometric Theory*, McGraw-Hill, New York, NY.

Reinartz, W.J., Haenlein, M. and Henseler, J. (2009), "An empirical comparison of the efficacy of covariance-based and variance-based SEM", *International Journal of Research in Marketing*, Vol. 26 No. 4, pp. 332-344.

Rigdon, E.E. (2012), "Rethinking partial least squares path modeling: in praise of simple methods", *Long Range Planning*, Vol. 45 Nos 5/6, pp. 341-358.

Rigdon, E.E. (2014), "Rethinking partial least squares path modeling: breaking chains and forging ahead", *Long Range Planning*, Vol. 47 No. 3, pp. 161-167.

Rigdon, E.E., Becker, J.-M., Rai, A., Ringle, C.M., Diamantopoulos, A., Karahanna, E., Straub, D.W. and Dijkstra, T.K. (2014), "Conflating antecedents and formative indicators: a comment on Aguirre-Urreta and Marakas", *Information Systems Research*, Vol. 25 No. 4, pp. 780-784.

Rindskopf, D. (1984), "Using phantom and imaginary latent variables to parameterize constraints in linear structural models", *Psychometrika*, Vol. 49 No. 1, pp. 37-47.

Ringle, C.M., Sarstedt, M. and Mooi, E.A. (2010a), "Response-based segmentation using finite mixture partial least squares: theoretical foundations and an application to American Customer Satisfaction Index data", *Annals of Information Systems*, Vol. 8 No. 1, pp. 19-49.

Ringle, C.M., Sarstedt, M. and Schlittgen, R. (2010b), "Finite mixture and genetic algorithm segmentation in partial least squares path modeling: identification of multiple segments in a complex path model", in Fink, A., Lausen, B., Seidel, W. and Ultsch, A. (Eds), *Advances in Data Analysis, Data Handling and Business Intelligence*, Springer, Berlin and Heidelberg, pp. 167-176.

Ringle, C.M., Sarstedt, M. and Schlittgen, R. (2014), "Genetic algorithm segmentation in partial least squares structural equation modeling", *OR Spectrum*, Vol. 36 No. 1, pp. 251-276.

Ringle, C.M., Sarstedt, M. and Straub, D.W. (2012), "Editor's comments: a critical look at the use of PLS-SEM in MIS Quarterly", *MIS Quarterly*, Vol. 36 No. 1, pp. iii-xiv.

Ringle, C.M., Wende, S. and Will, A. (2010c), "Finite mixture partial least squares analysis: methodology and numerical examples", in Esposito Vinzi, V., Chin, W.W., Henseler, J. and Wang, H. (Eds), *Handbook of Partial Least Squares: Concepts, Methods and Applications*, Springer, Berlin, pp. 195-218.

Rönkkö, M. and Evermann, J. (2013), "A critical examination of common beliefs about partial least squares path modeling", *Organizational Research Methods*, Vol. 16 No. 3, pp. 425-448.

Sahmer, K., Hanafi, M. and Qannari, M. (2006), "Assessing unidimensionality within the PLS path modeling framework", in Spiliopoulou, M., Kruse, R., Borgelt, C., Nürnberger, A. and Gaul, W. (Eds), *From Data and Information Analysis to Knowledge Engineering*, Springer, Berlin, pp. 222-229.

Sarstedt, M., Henseler, J. and Ringle, C. (2011), "Multi-group analysis in partial least squares (PLS) path modeling: alternative methods and empirical results", *Advances in International Marketing*, Vol. 22 No. 1, pp. 195-218.

Sarstedt, M., Ringle, C.M., Henseler, J. and Hair, J.F. (2014), "On the emancipation of PLS-SEM: a commentary on Rigdon (2012)", *Long Range Planning*, Vol. 47 No. 3, pp. 154-160.

Shmueli, G. and Koppius, O.R. (2013), "Predictive analytics in information systems research", *MIS Quarterly*, Vol. 35 No. 3, pp. 553-572.

Sijtsma, K. (2009), "On the use, the misuse, and the very limited usefulness of Cronbach's alpha", *Psychometrika*, Vol. 74 No. 1, pp. 107-120.

Streukens, S., Wetzels, M., Daryanto, A. and de Ruyter, K (2010), "Analyzing factorial data using PLS: application in an online complaining context", in Esposito Vinzi, V., Chin, W.W., Henseler, J. and Wang, H. (Eds), *Handbook of Partial Least Squares: Concepts, Methods and Applications*, Springer, Heidelberg, Dordrecht, London and New York, NY, pp. 567-587.

Tenenhaus, A. and Tenenhaus, M. (2011), "Regularized generalized canonical correlation analysis", *Psychometrika*, Vol. 76 No. 2, pp. 257-284.

Tenenhaus, M. (2008), "Component-based structural equation modelling", *Total Quality Management & Business Excellence*, Vol. 19 No. 7, pp. 871-886.

Tenenhaus, M., Amato, S. and Esposito Vinzi, V. (2004), "A global goodness-of-fit index for PLS structural equation modelling", *Proceedings of the XLII SIS Scientific Meeting, CLEUP, Padova*, pp. 739-742.

Tenenhaus, M., Esposito Vinzi, V., Chatelin, Y.-M. and Lauro, C. (2005), "PLS path modeling", *Computational Statistics & Data Analysis*, Vol. 48 No. 1, pp. 159-205.

Turkyilmaz, A., Oztekin, A., Zaim, S. and Fahrettin Demirel, O. (2013), "Universal structure modeling approach to customer satisfaction index", *Industrial Management & Data Systems*, Vol. 113 No. 7, pp. 932-949.

Voorhees, C.M., Brady, M.K., Calantone, R. and Ramirez, E. (forthcoming), "Discriminant validity testing in marketing: an analysis, causes for concern, and proposed remedies", *Journal of the Academy of Marketing Science* (in print).

Wold, H.O.A. (1974), "Causal flows with latent variables: partings of the ways in the light of NIPALS modelling", *European Economic Review*, Vol. 5 No. 1, pp. 67-86.

Wold, H.O.A. (1982), "Soft modeling: the basic design and some extensions", in Jöreskog, K.G. and Wold, H.O.A. (Eds), *Systems Under Indirect Observation*, North-Holland, Amsterdam, pp. 1-54.

Zhao, X., Lynch, J.G. and Chen, Q. (2010), "Reconsidering Baron and Kenny: myths and truths about mediation analysis", *Journal of Consumer Research*, Vol. 37 No. 2, pp. 197-206.

Ziggers, G.-W. and Henseler, J. (forthcoming), "The reinforcing effect of a firm's customer orientation and supply-base orientation on performance", *Industrial Marketing Management* (in print).

About the authors

Professor Jörg Henseler holds the Chair of Product-Market Relations at the University of Twente, The Netherlands. His research interests include structural equation modeling and the interface of marketing and design research. He has published in *Computational Statistics and Data Analysis*, *European Journal of Information Systems*, *European Journal of Marketing*, *International Journal of Research in Marketing*, *Journal of the Academy of Marketing Science*, *Journal of Service Management*, *Journal of Supply Chain Management*, *Long Range Planning*, *Management Decision*, *MIS Quarterly*, *Organizational Research Methods*, and *Structural Equation Modeling – An Interdisciplinary Journal*, among others. He has edited two handbooks and chaired two conferences on PLS path modeling. An author of the ADANCO computer program, he lectures worldwide on theory and applications of structural equation models. Professor Jörg Henseler is the corresponding author and can be contacted at: j.henseler@utwente.nl

Dr Geoffrey Hubona is the Founder of The Georgia R School. He has been conducting in-person and online workshops and classes on PLS and R for several years. He is an active Researcher and has been published in *MIS Quarterly, ACM Transactions on Computer-Human Interaction, IEEE Transactions on Systems, Man and Cybernetics, Data Base for Advances in Information Systems, Information & Management, Information Technology & People, Journal of Global Information Management, International Journal of Human Computer Studies, International Journal of Technology and Human Interaction, Journal of Organizational and End User Computing*, and *Journal of Information Technology Management*.

Pauline Ash Ray is an Associate Professor at the Thomas University, Thomasville, Georgia. Her research interests include management of change during implementation of information systems.