Industrial Marketing Management 128 (2025) A8–A15

Contents lists available at ScienceDirect

Industrial Marketing Management

journal homepage: www.elsevier.com/locate/indmarman

Why researchers should be cautious about using PLS-SEM[☆]

Jörg Henseler[a,d,*], Florian Schuberth[a], Nick Lee[b], Ildikó Kemény[c]

[a] Faculty of Engineering Technology, University of Twente, P.O. Box 217, 7500 AE Enschede, the Netherlands
[b] Warwick Business School, The University of Warwick, Coventry, CV4 7AL, UK
[c] Corvinus University of Budapest, Fövám tér 8, 1093 Budapest, Hungary
[d] Nova Information Management School, Universidade Nova de Lisboa, Portugal

ARTICLE INFO

Keywords:
PLS-SEM
Consistent PLS
Reflective measurement
Scenario analysis
Bias

ABSTRACT

A recent contribution to Industrial Marketing Management by Guenther, Guenther, Ringle, Zaefarian, and Cartwright (2023) aims at "[i]mproving PLS-SEM use." It assumes that the use of PLS-SEM could be improved by more strictly following existing PLS-SEM guidelines and using advanced techniques. Unfortunately, such "improved" use of PLS-SEM does not necessarily translate into more rigorous scientific analyses. In this commentary, we show that the PLS-SEM guidelines themselves are problematic and that using PLS-SEM can lead to erroneous conclusions. It is, therefore, not so much the use of PLS-SEM that needs improvement, but the literature on PLS-SEM itself, particularly its guidelines. The commentary concludes with recommendations for analysts and questions for PLS-SEM proponents to stimulate further research on PLS-SEM.

1. Introduction

An increasingly popular analytical method is partial least squares structural equation modeling (PLS-SEM, Guenther et al., 2023; Hair, Ringle, & Sarstedt, 2011; Sarstedt et al., 2022) – not only in B2B marketing research but also in business and social sciences generally.[1] Since the scientific enterprise can generally benefit from the correct use of research methods, Guenther et al. (2023) reviewed how PLS-SEM has been applied in B2B marketing research. Based on their findings, they proposed that researchers could improve the rigor of their PLS-SEM analyses if they (a) follow PLS-SEM guidelines for reflective and formative measurement models as presented by Hair, Hult, Ringle, and Sarstedt (2021), and (b) use advanced modeling approaches that build on PLS-SEM.

In principle, aiming to improve the use of a research method is a laudable endeavor. However, to improve the use of PLS-SEM in the way Guenther et al. (2023) suggested, the PLS-SEM results must be meaningful for reflective and formative measurement models so that it makes sense to build on them. Unfortunately, PLS-SEM is not suitable for reflective measurement models (e.g., Dijkstra & Henseler, 2015b; Hui & Wold, 1982; Sarstedt, Hair, Ringle, Thiele, & Gudergan, 2016; Schuberth et al., 2023b; Schuberth et al., 2023a). Similarly, the usefulness of existing PLS-SEM guidelines for assessing measurement models is questionable (e.g., Rönkkö et al., 2016; Rönkkö, Lee, Evermann, McIntosh, & Antonakis, 2023a; Schuberth, 2021). Consequently, following PLS-SEM guidelines that promote the use of PLS-SEM for reflective measurement models and the use of advanced modeling approaches that build on PLS-SEM results does not lead to more scientific rigor.

In our commentary, we reiterate, explain and demonstrate the core criticisms of PLS-SEM. Our goal is to enable researchers in B2B marketing and beyond to make an informed decision about which structural equation modeling method to use for their analysis. In addition, since the raised concerns have remained almost unheard in most of the PLS-

☆ Jörg Henseler gratefully acknowledges financial support from FCT Fundação para a Ciência e a Tecnologia (Portugal), national funding through a research grant from the Information Management Research Center – MagIC/NOVA IMS (UIDB/04152/2020). During the preparation of this work the authors used DeepL in order to improve the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication. The authors thank Alexandra Elbakyan for her efforts in making science accessible. Last but not least, the authors especially thank IMM's Associate Editor for Special Issues, Selma Kadić-Maglajlić, for granting the opportunity to write this commentary.

* Corresponding author.
E-mail addresses: j.henseler@utwente.nl (J. Henseler), f.schuberth@utwente.nl (F. Schuberth), nick.lee@wbs.ac.uk (N. Lee), ildiko.kemeny@uni-corvinus.hu (I. Kemény).

[1] We recognise an ongoing current debate on whether it is correct to refer to PLS-SEM as a structural equation modeling technique (cf. Rönkkö, McIntosh, Antonakis, & Edwards, 2016). We use the term PLS-SEM for consistency with the relevant literature only; thus, our use here should not be taken as an endorsement of the usefulness of the term.

https://doi.org/10.1016/j.indmarman.2024.01.017
Received 19 July 2023; Received in revised form 27 November 2023; Accepted 24 January 2024
Available online 26 May 2025

SEM guidelines, including the ones Guenther et al. (2023) put forward, we challenge future PLS-SEM research to take a stand on these criticisms, adjust the claims about PLS-SEM, and update the guidelines.

2. PLS-SEM is not suitable for estimating reflective measurement models

Arguably, one of the most important features of structural equation modeling (SEM) is the ability to specify measurement models. Measurement models are representations of auxiliary theories that postulate a certain link between a theoretical construct and one or more observed variables (Blalock, Hubert, & Costner, 1969; Henseler & Schuberth, 2021). They map the idea of measurement into a statistical model, which reflects a researcher's understanding of the world's mechanisms, namely of how unobserved theoretical constructs are related to observed variables.

Most studies using PLS-SEM indicate that they rely on reflective measurement models (cf. Hair, Hollingsworth, Randolph, & Chong, 2017; Magno, Cassia, & Ringle, 2024; Ringle, Sarstedt, Mitchell, & Gudergan, 2020), and Sarstedt et al. (2022, p. 14) "observe a clear shift toward reflective measurement in PLS-SEM studies." Reflective measurement assumes that the observed variables are caused by the theoretical construct. In the words of Guenther et al. (2023, p. 133), for "reflectively measured constructs, researchers assume a relationship from the theoretical construct to the indicators in its measurement model (e.g., the indicators reflect their construct)." Fornell, Rhee, and Yi (1991, pp. 316–317, typo corrected, equation numbers added) formalize reflective measurement as follows:

> If O is the observed measure, T, the true score and e an error component, it is well-known that the reflective specification is:

$$O = T + e \tag{1}$$

with the assumptions that

$$\mathrm{E}(e) = 0 \tag{2}$$

$$\mathrm{Cov}(T, e) = 0 \tag{3}$$

$$\mathrm{Cov}(e_i, e_j) = 0 \tag{4}$$

For a set of observed variables x_1, x_2, and x_3 conforming to a reflective measurement model, it would mean that each observed variable x_i can be explained by an underlying latent variable, ξ, and an individual error component, δ_i. A reflective measurement model with three variables, therefore, takes the shape of Fig. 1a.

In PLS-SEM, reflective measurement models are visualized as depicted in Fig. 1b (cf. Chin, 1998; Fornell & Bookstein, 1982; Tenenhaus, Vinzi, Chatelin, & Lauro, 2005). The PLS-SEM literature typically regards Fig. 1b as a valid representation of a reflective measurement model, and thus, as nothing but a simplified version of Fig. 1a (cf. Hair et al., 2021). As Hair et al. (2011, p. 141) put it: "PLS-SEM can handle […] reflective measurement models." Similar statements can also be found in more recent PLS-SEM literature (e.g., Sarstedt et al., 2022), and Guenther et al. (2023) regard the presence of constructs with reflective measurement models as a rationale for using PLS-SEM. However, as has been known for almost half a century, PLS-SEM does not provide consistent parameter estimates of reflective measurement models (e.g., Hui & Wold, 1982). Instead, PLS-SEM estimates the parameters of composite models (e.g., G. Cho, Sarstedt, & Hwang, 2022; Dijkstra, 2017).[2]

The composite model assumes that the theoretical construct is a linear combination of the observed variables (e.g., Hwang et al., 2021; Yu, Zaza, Schuberth, & Henseler, 2021). If as many non-redundant composites are created as there are observed variables, one can also express the observed variables in terms of the composites, i.e., one can invert the direction of the arrows (Henseler, 2021). This is done in the so-called (Henseler–Ogasawara (H–O) specification of composites (Schuberth, 2023; Yu, Schuberth, & Henseler, 2023). Fig. 1c shows such a specification of a composite model. To emphasize the nature of composites, they appear as hexagons. Although the arrows point from ξ to the observed variables, ξ is nevertheless a composite. One can easily verify that PLS-SEM actually estimates the model depicted in Fig. 1c: A linear regression model that regresses an observed variable (here: x_2) on the according construct scores and the residuals of all other observed variables explains all the variance in the observed variable. Thus, there is no residual. Moreover, the residuals in PLS-SEM are usually correlated.

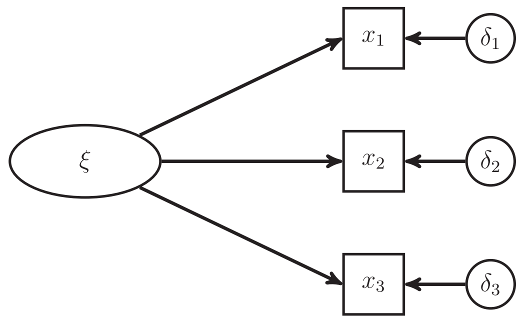
Considering that Fig. 1b only visualizes that part of the measurement model which reflective measurement models and composite models have in common, the visual representation of "reflective" measurement in PLS-SEM appears problematic: It conceals the fact that PLS-SEM usually does not satisfy Eqs. (1) and (4), and therefore is not suitable for estimating reflective measurement models.

Confounding composite models with reflective measurement models can have detrimental consequences for research. The composite model does not map the idea of reflective measurement (Henseler & Schuberth, 2023). This is also because measurement requires causal connections between the measures and their measurand (Cadogan & Lee, 2023b). If a composite is used as stand-in for a reflectively measured latent variable, the composite is contaminated by random measurement error. Consequently, PLS-SEM parameter estimates involving latent variables are biased due to attenuation (Cohen, Cohen, Teresi, Marchi, & Velez, 1990). This issue is well known in the methodological literature on PLS-SEM (e.g., Dijkstra, 1985; Hui & Wold, 1982; Schuberth, Schamberger, & Henseler, 2024; Schuberth, Zaza, & Henseler, 2023d) and widely discussed in the literature on error-in-variables (e.g., Buonaccorsi, 2010; Carroll, Ruppert, Stefanski, & Crainiceanu, 2006), a strand of literature that focuses on the consequences of variables contaminated by error.

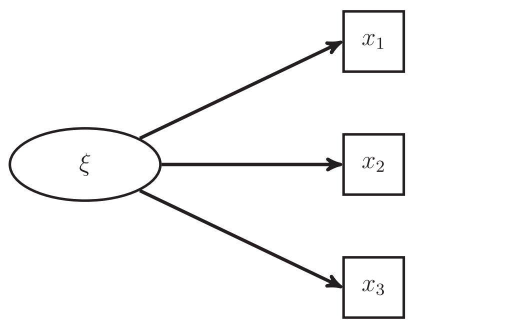
To illustrate how PLS-SEM can lead researchers to erroneous conclusions if reflectively measured latent variables are involved, we make use of a hypothetical B2B marketing study that aims to examine the effect a novel marketing tactic has on buyers' repurchase intention. The empirical data consists of 100 observations and five variables: TACTIC is a dichotomous variable that flags whether a specific marketing tactic was used (value of 1), or not (value of 0). The variable TRUST captures the true level of a buyer's trust. The two variables TR_IND1 and TR_IND2 are responses to survey questions about the buyer's trust measured on 7-point scales. They are only caused by TRUST, with the standardized effects of TRUST on the two variables being 0.70 and 0.80, respectively. Finally, the variable INTENT denotes the buyer's intention to repurchase a product; it has discrete values ranging from 0 to 100. The standardized regression coefficients of TRUST and TACTIC on Repurchase Intention are 0.81 and 0.15, respectively. This means that the marketing tactic has a positive effect on the buyer's intention to repurchase a product (ceteris paribus). There is some degree of multicollinearity between TRUST and TACTIC, i.e., a correlation of −0.76. However, the variance inflation factors of 2.37 are clearly below conventional thresholds (cf. Hair, Black, Babin, & Anderson, 2019a). The relationships between the variables are illustrated in Fig. 2[3] As PLS-SEM does not allow the inclusion of observed variables directly in the structural model, each observed variable is modeled as a single-indicator construct (e.g., Benitez, Henseler,

[2] PLS-SEM using Mode A estimates composite models such that the weight parameters are proportional to the loadings of a common factor model (Henseler & Schuberth, 2023).

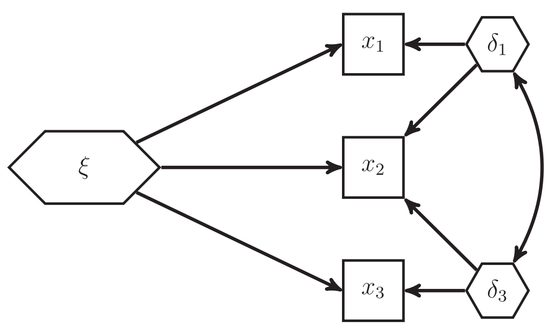
[3] Note that all values are rounded to two digits. The data and R code we used can be accessed via https://osf.io/q98zc/?view_only=efa82d3c92514d54871834b9a2c43f53.



(a) A reflective measurement model with three observed variables (what PLS-SEM allegedly does)



(b) How PLS-SEM literature and software symbolizes a reflective measurement model with three indicators



(c) A composite model with three indicators and three composites (what PLS-SEM actually does)

Fig. 1. When PLS-SEM purports to estimate reflective measurement models, it actually estimates composite models.

Castillo, & Schuberth, 2020).

In a real research situation, though, the researcher would not have access to the values of TRUST, but only to the survey data of TR_IND1 and TR_IND2, i.e., the two variables become reflective measures of TRUST. Fig. 3 shows the results for a researcher who applies PLS-SEM using Mode A to estimate this structural equation model. Apparently, the effects of the latent variable on both indicators are inflated, which makes them look even more reliable than they really are. Cronbach's α has a value of 0.71, and composite reliability ρ_C is 0.87, i.e., ρ_C is substantially upward-biased.[4] This is not surprising, because it is calculated on the basis of the upward-biased PLS-SEM indicator loading estimates (Gefen, Straub, & Rigdon, 2011). A PLS-SEM user would not have any reason for supposing the reliability to be insufficient, especially because PLS-SEM guidelines (Hair et al., 2021, p. 119) regard ρ_C as "technically more appropriate" than Cronbach's α, and Sarstedt et al. (2022, p. 1050) state that "when Cronbach's alpha does not raise concerns in a PLS-SEM analysis, the construct measurement can a fortiori be expected to exhibit sufficient levels of internal consistency reliability." Moreover, the path coefficients provided by PLS-SEM are heavily distorted. The effect of trust on intention is underestimated by 38%, and – most disturbingly – the effect of the marketing tactic on the buyer's intention to repurchase a product has the opposite sign. Thus, although the actual effect of the marketing tactic is positive, the researcher using PLS-SEM would observe a negative effect. Hence, if PLS-SEM were employed in this B2B marketing study, it would have led to erroneous conclusions, and the scientific community would be denied a promising new marketing tactic.

To verify that the erroneous conclusions are not a common problem of SEM but purely attributable to PLS-SEM, one could use consistent SEM estimators such as covariance-based SEM with maximum likelihood or consistent partial least squares (PLSc, Dijkstra & Henseler, 2015a,b) to estimate the parameters of our illustrative example. Fig. 4 shows the results obtained by PLSc isis. Note that values are rounded to two digits. As one can clearly see, all relationships are in line with the case where Trust is directly observed (see Fig. 2).

Conclusion: Much of the PLS-SEM literature describes and visualizes structural equation models in a way that gives the false impression that

[4] The actual value of ρ_C is 0.72.

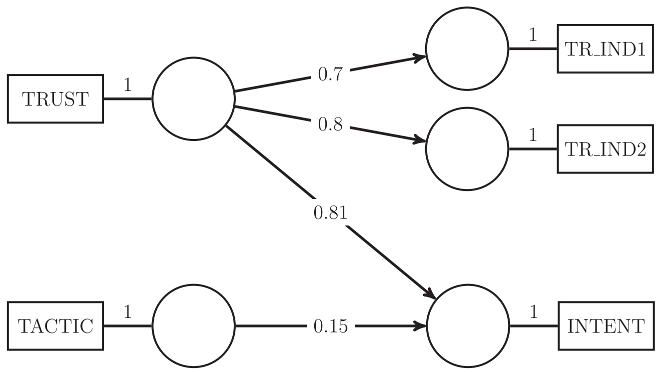


Fig. 2. Relationships estimated by PLS-SEM when all variables are observed.

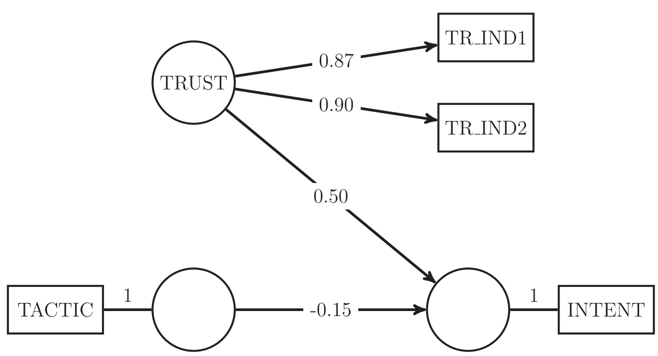


Fig. 3. Relationships estimated by PLS-SEM.

PLS-SEM can deal with reflective measurement of latent variables. However, PLS-SEM is not suitable for estimating the parameters of this type of measurement model. PLS-SEM parameter estimates of reflective measurement models will be biased and the results can be misleading.

3. Extant PLS-SEM guidelines for model assessment are problematic

MacCallum (2003, p. 136) reminds us that "our models can be useful if they are not grossly wrong – useful for prediction, for testing and developing theories, for clarifying the nature of the world." Therefore, model assessment is not only a crucial element of SEM in general, but also an important part of Guenther et al.'s (2023) endeavor to improve the use of PLS-SEM. Concretely, Guenther et al. (2023) recommend that users follow extant PLS-SEM guidelines for model assessment more strictly, and particularly, those Hair and colleagues (cf. Hair et al., 2021; Hair, Risher, Sarstedt, & Ringle, 2019b; Hair, Sarstedt, Pieper, & Ringle, 2012) proposed. Implicitly, this recommendation is built on the premise that the extant guidelines are useful and contribute to the well-being of science. Unfortunately, the PLS-SEM guidelines for reflective measurement model assessment are problematic on several counts (e.g., Hubona, Schuberth, & Henseler, 2021; Rönkkö et al., 2016; Rönkkö & Evermann, 2013; Schuberth, 2021; Schuberth, Hubona, et al., 2023a). First, while PLS-SEM estimates composite models (Sabol, Hair, Cepeda, & Roldàn, 2023), the proposed assessment criteria are not designed for composite models. Thus, their interpretation for composite models is at best unclear, and at worst misleading. In fact, these assessment criteria were developed in the context of common factor models for which PLS-SEM produces biased estimates. Therefore, second, the PLS-SEM assessment criteria for reflective measurement models are misleading if they are based on biased PLS-SEM common factor model parameter estimates.

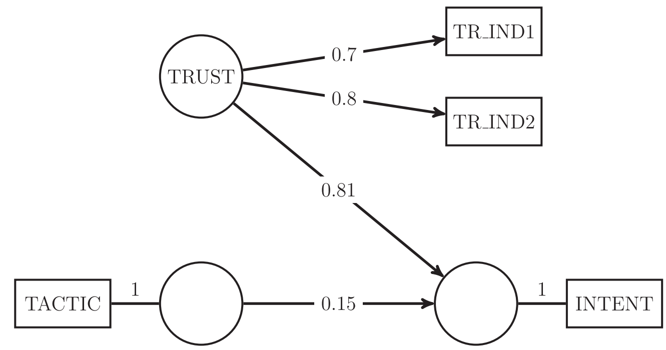


Fig. 4. Relationships estimated by a consistent estimator (here: results obtained through PLSc).

We will elaborate on these issues in the following subsections.

3.1. *Assessment criteria for reflective measurement models have not been designed for composite models*

To assess reflective measurement models, Guenther et al. (2023) recommend that users consider various criteria. As shown in Table 1, these criteria include indicator reliability, composite reliability, average variance extracted (AVE, Fornell & Larcker, 1981), and the heterotrait-monotrait ratio of correlations (HTMT, Henseler et al., 2015). Notably, all these criteria have been developed against the backdrop of the common factor model or the true score theory, the latter being a special case of the common factor model (e.g., McDonald, 1999), which assumes that an indicator can be decomposed into a true score and an error score (e.g., Lord & Novick, 2008). In contrast, and as Guenther et al. (2023) emphasized, PLS-SEM does not consistently estimate common factor models; rather, it estimates composite models. In this case, the use of these assessment criteria represents a mismatch between the criteria's assumptions and the model estimated by PLS-SEM. Consequently, the assessment criteria's behavior and interpretation are not clear in this situation and their results can be misleading, as we will show next.

Indicator reliability quantifies the share of variance in an indicator that can be explained by its underlying common factor (Jöreskog, 1971). Its interpretation in the context of a composite model is unclear, because the composite model – in contrast to the common factor model – does not model an observed variable as one contaminated with measurement error.

"To assess internal consistency reliability (i.e., the extent to which the set of indicators consistently reflect the underlying construct), researchers should report composite reliability ρ_A. Additional optional criteria are Cronbach's α and the composite reliability ρ_C" (Guenther et al., 2023, p. 134). Reliability coefficients such as Cronbach's α, composite reliability ρ_C, also known as Jöreskog ρ (Jöreskog, 1971), and composite reliability ρ_A, also known as Dijkstra-Henseler's ρ_A (Dijkstra & Henseler, 2015b), quantify the share of true score variance in a linear combination of indicators. Specifically, it is defined as the variation in a linear combination of indicators that can be explained by an underlying common factor divided by the total variation of that linear combination (E. Cho, 2016). Hereby, each coefficient makes different assumptions. While both Cronbach's α and composite reliability ρ_C assume that the linear combination of the indicators is obtained as a simple sum, Cronbach's alpha makes an additional, more restrictive, assumption about the indicators. It assumes an (essentially) tau-equivalent measurement model, which is a restrictive form of the common factor model that entails all unstandardized indicator factor loadings are equal (see Novick & Lewis, 1967). In contrast, and similar to the composite reliability ρ_C, Dijkstra-Henseler's ρ_A is less restrictive and assumes a congeneric

Table 1
PLS-SEM assessment criteria that rely on a common factor model, although PLS-SEM does not estimate this model.

Criterion	Formula	Explanation	Reference
Indicator reliability	$\frac{\widehat{\lambda}_i^2}{\widehat{\lambda}_i^2 + \widehat{var(\varepsilon_i)}}$	Estimates the variance in an indicator that can be explained by its underlying common factor.	Werts, Linn, & Jöreskog (1974) Jöreskog (1971)
Cronbach's α	$\frac{K^2 \bar{s}_{ij}}{\widehat{\sigma}_Y^2}$	Estimates the reliability of a sum of (essentially) tau-equivalent measures, i.e., the share of variance in that sum which can be explained by the underlying common factor.	Guttman (1945) Cronbach (1951) Falk and Savalei (2011)
Composite reliability ρ_C	$\frac{\left(\sum_{i=1}^{K}\widehat{\lambda}_i\right)^2}{\left(\sum_{i=1}^{K}\widehat{\lambda}_i\right)^2 + \left(\sum_{i=1}^{K}\widehat{var(\varepsilon_i)}\right)^2}$	Estimates the reliability of a sum of congeneric measures, i.e., the share of variance in that sum which can be explained by the underlying common factor.	Werts, Linn, & Jöreskog (1974) Jöreskog (1971)
Composite reliability ρ_A	$(\widehat{\boldsymbol{w}}'\widehat{\boldsymbol{w}})^2 \frac{\widehat{\boldsymbol{w}}'(\boldsymbol{S} - diag(\boldsymbol{S}))\widehat{\boldsymbol{w}}}{\widehat{\boldsymbol{w}}'(\widehat{\boldsymbol{w}}\widehat{\boldsymbol{w}}' - diag(\widehat{\boldsymbol{w}}\widehat{\boldsymbol{w}}'))\widehat{\boldsymbol{w}}}$	Estimates the reliability of a weighted sum of congeneric measures using PLS-SEM Mode A weights, i.e., the share of variance in that weighted sum which can be explained by the underlying common factor.	Dijkstra & Henseler (2015a)
Heterotrait-monotrait ratio of correlations	$\frac{1}{(K_iK_j)} \frac{\sum_{g=1}^{K_i}\sum_{h=1}^{K_j} r_{ig,jh}}{\sqrt{\frac{2}{K_i(K_i-1)}\sum_{g=1}^{K_i-1}\sum_{h=g+1}^{K_i} r_{ig,ih}\cdot\frac{2}{K_j(K_j-1)}\sum_{g=1}^{K_j-1}\sum_{h=g+1}^{K_j} r_{jg,jh}}}$	Estimates the correlation between two common factors with tau-equivalent measures.	Henseler, Ringle, & Sarstedt (2015)
Average variance extracted	$\frac{\sum_{i=1}^{K}\widehat{\lambda}_i^2}{\sum_{i=1}^{K}\widehat{\lambda}_i^2 + \sum_{i=1}^{K}\widehat{var(\varepsilon_i)}}$	Estimates the amount of variance in a set of measures that can be explained by its underlying common factor.	Fornell & Larcker (1981)

$\widehat{\lambda}_i$: estimated factor loading of indicator x_i; $\widehat{var(\varepsilon_i)}$: estimated variance of the error term that accounts for the variance in indicator x_i which cannot be explained by the underlying common factor ξ; $\bar{s}_{ij}$: average sample covariance of the indicators belonging to a common factor; K: number of indicators belonging to a common factor ξ; $\widehat{\sigma}_Y^2$: sample variance of the sum of the indicators belonging to a common factor; $\widehat{\boldsymbol{w}}$: PLS-SEM Mode A weight vector. $\boldsymbol{S}$: sample correlation matrix of the indicators belonging to a common factor; $r_{ig,jh}$ sample correlation between the indicators x_g and x_h that belong to common factor ξ_i and ξ_j, respectively; K_i,K_j: number of indicators belonging to the i-th and j-th common factor, respectively.

measurement model, i.e., a common factor model in which all parameters are allowed to be different (E. Cho, 2016; Dijkstra & Henseler, 2015a). In addition, ρ_A assumes that PLS-SEM Mode A weights were used for calculating the linear combination of the indicators. Since the composite model builds on composites and not on common factors, there is an obvious conflict with the assumptions of the three reliability coefficients, and consequently, their interpretation for composite models is not clear.

To assess discriminant validity, Guenther et al. (2023) propose that the HTMT be considered. The HTMT is a consistent estimator for the correlation between two common factors with tau-equivalent measures (Henseler et al., 2015). To relax this rigid assumption, the HTMT2 which estimates the correlation between two common factors with congeneric measures, was introduced (Roemer, Schuberth, & Henseler, 2021). Both criteria exploit the covariance structure of the indicators implied by the common factor model (Henseler et al., 2015; Roemer et al., 2021). Hence, the use of these criteria in the context of composite models is questionable because "[c]omposite indicators for a [composite] have no specific covariance pattern since their covariation is not caused by the construct" (Hwang, Sarstedt, Cho, Choo, & Ringle, 2023, p. 264). To illustrate this issue, we apply the HTMT to a composite population model. The population consists of two composites, ξ_1 and ξ_2, each formed by three indicators using correlation weights. The indicators are standardized and the correlations between the indicators of one block equal 0.5. The resulting composite loadings are all equal to 0.82. The correlation between the two composites is set to 0.7. Fig. 5 shows the population model. PLS-SEM using Mode A is able to correctly retrieve the composite loadings and the correlation between the two composites. However, applying the HTMT to this population model leads to a value of 0.93.[5] Obviously, this value differs from 0.7, i.e., from the true

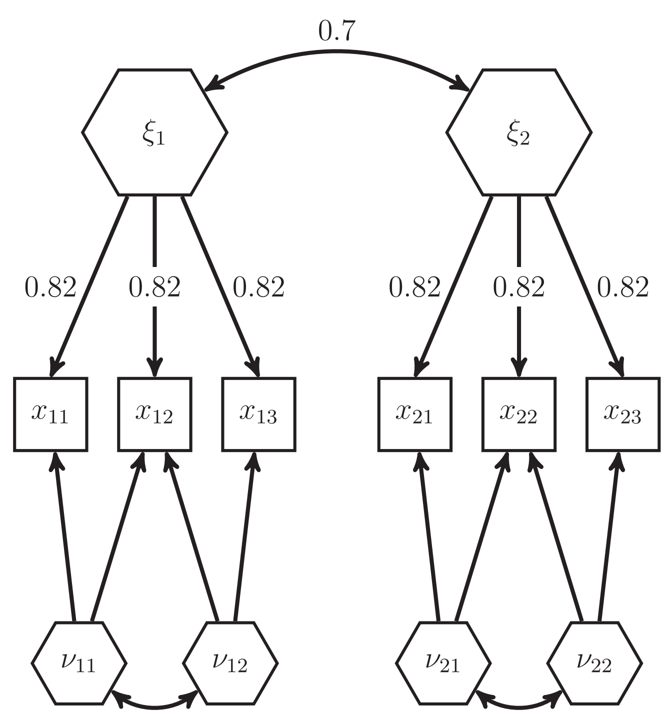


Fig. 5. Population model with two composites formed by correlation weights.

correlation between the two composites, which highlights our concerns about the usefulness of the HTMT and HTMT2 for composite models.

Finally, the AVE is computed based on common factor model parameter estimates (Rigdon, 2013). It quantifies "the amount of variance that is captured by the construct in relation to the amount of

[5] The data we used and R code can be accessed via https://osf.io/q98zc/?view_only=efa82d3c92514d54871834b9a2c43f53.

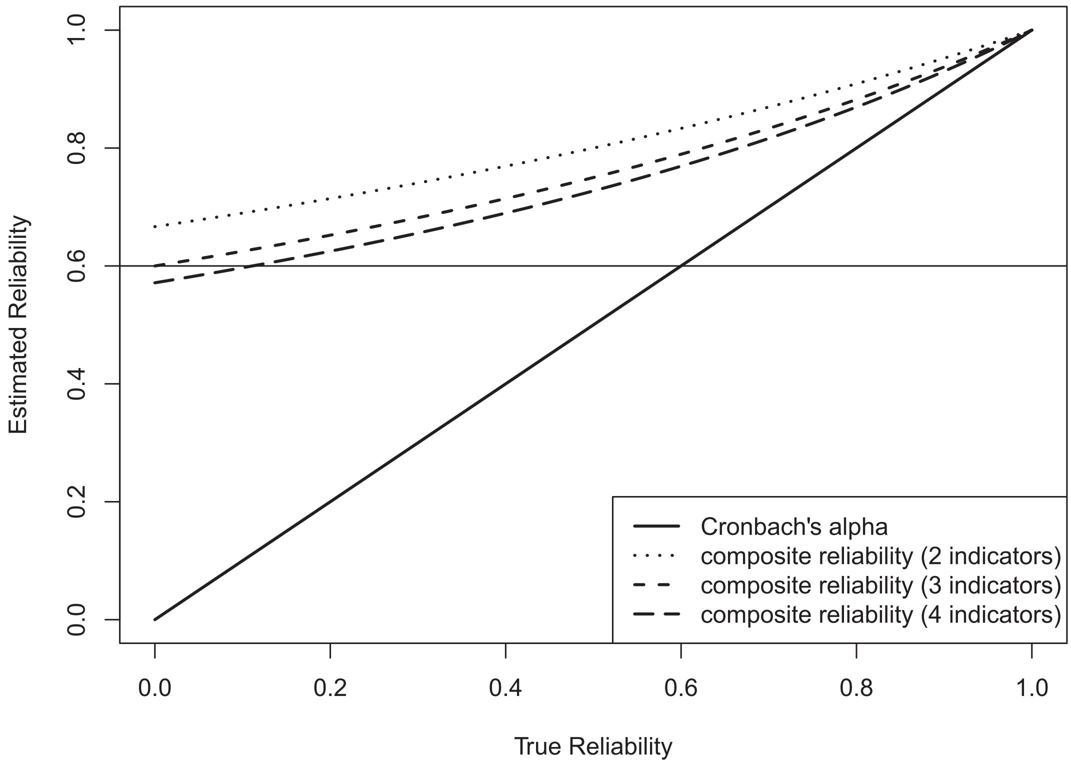


Fig. 6. Relationship between true reliability and that estimated by PLS-SEM.

variance due to measurement error" (Fornell & Larcker, 1981, p. 45). Since the composite model does not distinguish between variance that is explained by the construct and variance that is due to measurement error, the interpretation of the AVE in this context is unclear.

3.2. *Assessment criteria are biased if based on PLS-SEM parameter estimates for common factor models*

All of the proposed model assessment criteria for reflective measurement models have been developed against the backdrop of a common factor model. Since PLS-SEM produces biased estimates for common factor model parameters, assessment criteria like indicator reliability, composite reliability ρ_C, and AVE will be biased, too. Although some assessment criteria can be calculated without estimating a common factor model, such as Cronbach's α, composite reliability ρ_A, and the HTMT/HTMT2, rules for their assessment are directly based on the assumption that the common factor model holds. Consequently, researchers using PLS-SEM in combination with these model assessment criteria will most likely be misguided.

Indicator reliability quantifies the share of variance in an indicator that is explained by the underlying common factor. Since in PLS-SEM the factor loading estimates are upward-biased (Gefen et al., 2011), also the indicator reliability will be upward biased.

Similar to the indicator reliability, the composite reliability ρ_C is also based on the common factor model parameters. As Guenther et al. (2023, p. 134) acknowledged, the "composite reliability ρ_C [is] too liberal, i.e., upward biased". In the PLS-SEM context, the composite reliability ρ_C is upward-biased because it relies on loading estimates obtained through PLS-SEM Mode A, which are upward-biased themselves. It should be noted that ρ_C is generally not upward-biased (cf. E. Cho, 2016). Fig. 6 gives an impression of ρ_C's upward-bias in the PLS-SEM context and shows that it is hardly possible to ever obtain values for ρ_C smaller than 0.60 if PLS-SEM estimates are used. This is particularly problematic because current PLS-SEM guidelines highlight that "values below 0.60 indicate a lack of internal consistency reliability" (Hair et al., 2021, p. 119). As a result, PLS-SEM practitioners following this guideline are unlikely to be warned of a lack of internal consistency reliability, regardless of the true reliability.

Finally, for calculating the AVE, the common factor loading estimates are required. Since PLS-SEM provides biased estimates of the factor loadings, the AVE will be biased, too. Consequently, the use of the AVE in the PLS-SEM context gives a distorted picture of the variance in a set of reflective measures which can be explained by the underlying common factor. Guenther et al. (2023, p. 134) also acknowledge this in stating "that AVE is not useful when the number of a construct's indicators is as small as two, in which case AVEs larger than 0.50 would always be obtained in PLS-SEM." This is due to the fact that biased PLS-SEM common factor loading estimates are used for the calculation.

Conclusion: The PLS-SEM assessment criteria for reflective measurement models are either used to assess a model for which the criteria were not designed, i.e., a composite model, which renders their interpretation questionable, or they are used to assess common factor models, for which PLS-SEM produces biased parameter estimates. The latter, in turn, renders most of the assessment criteria biased. Consequently, the use of the PLS-SEM assessment criteria cannot be recommended.

4. Implications

The recent IMM paper by Guenther et al. (2023) critically reviews how PLS-SEM has been applied in B2B marketing research, and identifies a number of shortcomings in the current use of PLS-SEM. Further, it presents reasons and motives for using PLS-SEM, reiterates PLS-SEM guidelines, and promotes the use of advanced PLS-SEM procedures. Based on Guenther et al. (2023), one could conclude that if anything were wrong with PLS-SEM, then it would be the way in which it is currently used. However, much deeper problems can be found in the methodological literature on PLS-SEM.

The vast majority of empirical studies using PLS-SEM contain reflectively measured constructs, and Guenther et al. (2023) suggest that

PLS-SEM can be used in such situations. As shown by mathematical proofs, various simulation studies and scenario analyses, including the one provided in our commentary above, this is not true. The use of PLS-SEM can lead to severely biased estimates and questionable conclusions drawn from the estimated model. Those cases in which advanced PLS-SEM procedures rely on the results of models with reflective measurement models – one could think of higher-order constructs, mediation analysis, or unobserved heterogeneity – are ones for which we do not share Guenther et al.'s optimistic view that advanced PLS-SEM procedures can "generate more robust findings, a deeper understanding, and novel insights." Consequently, we cannot recommend using PLS-SEM, including its advanced modeling approaches, in the way Guenther et al. (2023) promote. Instead, analysts could rather rely on updated guidelines that come without the identified problems (e.g., Benitez et al., 2020; Evermann & Rönkkö, 2023; Henseler, Hubona, & Ray, 2016).

Methodological research gives various reasons for being critical of PLS-SEM, particularly in the case of reflective measurement. However, analysis of the citations shows that PLS-SEM users are either unaware of such research, or routinely ignore it. This paper joins a number of other recent studies (e.g., Cadogan & Lee, 2023a, 2023b; Henseler & Schuberth, 2023; Rönkkö et al., 2023a; Rönkkö, Lee, Evermann, McIntosh, & Antonakis, 2023b; Schuberth, 2021; Henseler et al., 2024) in discussing critiques of PLS-SEM in more applied business and management research literature, in the hope of widening appreciation of these issues. We are encouraged by this growing trend, but we do notice that there are significant areas where more discussion is needed. For example, Guenther et al. (2023,Footnote 4) note that "the distinction between common factor and composite model is not equivalent to reflective and formative measurement – as sometimes suggested by PLS-SEM critics." Interestingly, they do not cite any of these 'critics', and therefore, we would ask Guenther et al. (2023), as well as other PLS-SEM proponents, to clarify the following questions: 1) What is the difference between a reflective measurement model and a common factor model, if any? 2) What does the data generating process look like in the case of reflective measurement? 3) How does a composite population model, i.e., the data generating process, map the idea of reflective measurement? In addition, PLS-SEM research still owes us proof of the efficacy of its assessment criteria. We hope that these questions stimulate a discussion on the narrative and the guidelines of PLS-SEM, which should be reconsidered.

CRediT authorship contribution statement

Jörg Henseler: Conceptualization, Funding acquisition, Writing – original draft. **Florian Schuberth:** Conceptualization, Formal analysis, Visualization, Writing – original draft. **Nick Lee:** Conceptualization, Writing – review & editing. **Ildikó Kemény:** Writing – review & editing.

Data availability

Data is made available via OSF: https://osf.io/q98zc/?view_only=efa82d3c92514d54871834b9a2c43f53.

References

Benitez, J., Henseler, J., Castillo, A., & Schuberth, F. (2020). How to perform and report an impactful analysis using partial least squares: Guidelines for confirmatory and explanatory IS research. *Information & Management, 2*(57), Article 103168. https://doi.org/10.1016/j.im.2019.05.003

Blalock, J., Hubert, M., & Costner, H. L. (1969). Multiple indicators and the causal approach to measurement error. *American Journal of Sociology, 75*(2), 264–273. https://doi.org/10.1086/224771

Buonaccorsi, J. (2010). *Measurement error: Models, methods, and applications*. Boca Raton, FL: CRC Press.

Cadogan, J. W., & Lee, N. (2023a). A miracle of measurement or accidental constructivism? How PLS subverts the realist search for truth. *European Journal of Marketing, 57*(6), 1703–1724. https://doi.org/10.1108/EJM-08-2020-0637

Cadogan, J. W., & Lee, N. (2023b). Scientific realism, the necessity of causal contact in measurement and emergent variables. *European Journal of Marketing, 57*(6), 1758–1779. https://doi.org/10.1108/ejm-06-2021-0454

Carroll, R. J., Ruppert, D., Stefanski, L. A., & Crainiceanu, C. M. (2006). *Measurement error in nonlinear models : A modern perspective*. Boca Raton, FL: Chapman & Hall/CRC.

Chin, W. W. (1998). The partial least squares approach for structural equation modeling. In G. Marcoulides (Ed.), *Modern methods for business research* (pp. 295–336). New York: Psychology Press.

Cho, E. (2016). Making reliability reliable: A systematic approach to reliability coefficients. *Organizational Research Methods, 19*(4), 651–682. https://doi.org/10.1177/1094428116656239

Cho, G., Sarstedt, M., & Hwang, H. (2022). A comparative evaluation of factor-and component-based structural equation modelling approaches under (in)correct construct representations. *British Journal of Mathematical and Statistical Psychology, 75*(2), 220–251. https://doi.org/10.1111/bmsp.12255

Cohen, P., Cohen, J., Teresi, J., Marchi, M., & Velez, C. N. (1990). Problems in the measurement of latent variables in structural equations causal models. *Applied Psychological Measurement, 14*(2), 183–196. https://doi.org/10.1177/014662169001400207

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334. https://doi.org/10.1007/bf02310555

Dijkstra, T. K. (1985). *Latent variables in linear stochastic models: Reflections on "maximum likelihood" and "partial least squares" methods* (Vol. 1). Amsterdam: Sociometric Research Foundation.

Dijkstra, T. K. (2017). A perfect match between a model and a mode. In H. Latan, & R. Noonan (Eds.), *Partial least squares path modeling* (pp. 55–80). Cham: Springer. https://doi.org/10.1007/978-3-319-64069-34.

Dijkstra, T. K., & Henseler, J. (2015a). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics & Data Analysis, 81*, 10–23. https://doi.org/10.1016/j.csda.2014.07.008

Dijkstra, T. K., & Henseler, J. (2015b). Consistent partial least squares path modeling. *MIS Quarterly, 39*(2), 297–316. https://doi.org/10.25300/MISQ/2015/39.2.02

Evermann, J., & Rönkkö, M. (2023). Recent developments in PLS. *Communications of the Association for Information Systems, 52*, 663–667. https://doi.org/10.17705/1CAIS.05229

Falk, C. F., & Savalei, V. (2011). The relationship between unstandardized and standardized alpha, true reliability, and the underlying measurement model. *Journal of Personality Assessment, 93*(5), 445–453. https://doi.org/10.1080/00223891.2011.594129

Fornell, C., & 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. https://doi.org/10.1177/002224378201900406

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.1177/002224378101800104

Fornell, C., Rhee, B.-D., & Yi, Y. (1991). Direct regression, reverse regression, and covariance structure analysis. *Marketing Letters, 2*(3), 309–320. https://doi.org/10.1007/BF00554134

Gefen, D., Straub, D. W., & Rigdon, E. E. (2011). An update and extension to SEM guidelines for administrative and social science research. *MIS Quarterly, 35*(2), iii–xiv.

Guenther, P., Guenther, M., Ringle, C. M., Zaefarian, G., & Cartwright, S. (2023). Improving PLS-SEM use for business marketing research. *Industrial Marketing Management, 111*(May), 127–142. https://doi.org/10.1016/j.indmarman.2023.03.010

Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika, 10*(4), 255–282. https://doi.org/10.1007/bf02288892

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019a). *Multivariate data analysis* (8th ed.). Cengage.

Hair, J. F., Hollingsworth, C. L., Randolph, A. B., & Chong, A. Y. L. (2017). An updated and expanded assessment of PLS-SEM in information systems research. *Industrial Management & Data Systems, 117*(3), 442–458. https://doi.org/10.1108/IMDS-04-2016-0130

Hair, J. F., Hult, G., Ringle, C. M., & Sarstedt, M. (2021). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). Sage.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice, 19*(2), 139–152. https://doi.org/10.2753/MTP1069-6679190202

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019b). When to use and how to report the results of PLS-SEM. *European Business Review, 31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Hair, J. F., Sarstedt, M., Pieper, T. M., & Ringle, C. M. (2012). 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. https://doi.org/10.1016/j.lrp.2012.09.008

Henseler, J. (2021). *Composite-based structural equation modeling: Analyzing latent and emergent variables*. New York, NY: Guilford Press.

Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: Updated guidelines. *Industrial Management & Data Systems, 116*(1), 2–20. https://doi.org/10.1108/IMDS-09-2015-0382

Henseler, J., Lee, N., Roemer, E., Kemény, I., Dirsehan, T., & Cadogan, J. W. (2024). Beware of the Woozle effect and belief perseverance in the PLS-SEM literature! *Electronic Commerce Research*. https://doi.org/10.1007/s10660-024-09849-y. in print.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*, 115–135. https://doi.org/10.1007/s11747-014-0403-8

Henseler, J., & Schuberth, F. (2021). Auxiliary theories. In *Composite-based Structural Equation Modeling: Analyzing Latent and Emergent Variables* (pp. 25–37). New York: Guilford Press.
Henseler, J., & Schuberth, F. (2023). Partial least squares as a tool for scientific inquiry: Comments on Cadogan and Lee. *European Journal of Marketing, 57*(6), 1737–1757. https://doi.org/10.1108/EJM-06-2021-0416
Hubona, G. S., Schuberth, F., & Henseler, J. (2021). A clarification of confirmatory composite analysis (CCA). *International Journal of Information Management, 61*, Article 102399. https://doi.org/10.1016/j.ijinfomgt.2021.102399
Hui, B. S., & Wold, H. (1982). Consistency and consistency at large of partial least squares estimates. In *Systems under indirect observation, Part II* (pp. 119–130).
Hwang, H., Cho, G., Jung, K., Falk, C., Flake, J., Jin, M. J., & Lee, S. H. (2021). An approach to structural equation modeling with both factors and components: Integrated generalized structured component analysis. *Psychological Methods, 26*(3), 273–294. https://doi.org/10.1037/met0000336
Hwang, H., Sarstedt, M., Cho, G., Choo, H., & Ringle, C. M. (2023). A primer on integrated generalized structured component analysis. *European Business Review*. https://doi.org/10.1108/ebr-11-2022-0224
Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika, 36* (2), 109–133. https://doi.org/10.1007/bf02291393
Lord, F. M., & Novick, M. R. (2008). *Statistical theories of mental test scores*. Information Age Publishing.
MacCallum, R. C. (2003). 2001 presidential address: Working with imperfect models. *Multivariate Behavioral Research, 38*(1), 113–139. https://doi.org/10.1207/S15327906MBR3801_5
Magno, F., Cassia, F., & Ringle, C. M. (2024). A brief review of partial least squares structural equation modeling (PLS-SEM) use in quality management studies. *The TQM Journal, 36*(5), 1242–1251. https://doi.org/10.1108/TQM-06-2022-0197
McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
Novick, M., & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurements. *Psychometrika, 32*(1), 1–13. https://doi.org/10.1007/BF02289400
Rigdon, E. E. (2013). Lee, Cadogan, and Chamberlain: an excellent point . . . but what about that iceberg? *AMS Review, 3*(1), 24–29. https://doi.org/10.1007/s13162-013-0034-0
Ringle, C. M., Sarstedt, M., Mitchell, R., & Gudergan, S. P. (2020). Partial least squares structural equation modeling in HRM research. *International Journal of Human Resource Management, 31*(12), 1617–1643. https://doi.org/10.1080/09585192.2017.1416655
Roemer, E., Schuberth, F., & Henseler, J. (2021). HTMT2–an improved criterion for assessing discriminant validity in structural equation modeling. *Industrial Management & Data Systems, 121*(12), 2637–2650. https://doi.org/10.1108/IMDS-02-2021-0082
Rönkkö, M., & Evermann, J. (2013). A critical examination of common beliefs about partial least squares path modeling. *Organizational Research Methods, 16*(3), 425–448. https://doi.org/10.1177/1094428112474693
Rönkkö, M., Lee, N., Evermann, J., McIntosh, C., & Antonakis, J. (2023a). Marketing or methodology? Exposing the fallacies of PLS with simple demonstrations. *European Journal of Marketing, 57*(6), 1597–1617. https://doi.org/10.1108/EJM-02-2021-0099
Rönkkö, M., Lee, N., Evermann, J., McIntosh, C. N., & Antonakis, J. (2023b). Rejoinder: Fractures in the edifice of PLS. *European Journal of Marketing, 57*(6), 1626–1640. https://doi.org/10.1108/EJM-07-2022-0508
Rönkkö, M., McIntosh, C. N., Antonakis, J., & Edwards, J. R. (2016). Partial least squares path modeling: Time for some serious second thoughts. *Journal of Operations Management, 47*, 9–27. https://doi.org/10.1016/j.jom.2016.05.002
Sabol, M. A., Hair, J. F., Cepeda, G., & Roldàn, J. L. (2023). PLS-SEM in information systems: Seizing the opportunity and marching ahead full speed. *Industrial Management and Data Systems, 123*(12), 2997–3017. https://doi.org/10.1108/IMDS-07-2023-0429
Sarstedt, M., Hair, J. F., Pick, M., Liengaard, B. D., Radomir, L., & Ringle, C. M. (2022). Progress in partial least squares structural equation modeling use in marketing research in the last decade. *Psychology & Marketing, 39*(5), 1035–1064. https://doi.org/10.1002/mar.21640
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. https://doi.org/10.1016/j.jbusres.2016.06.007
Schuberth, F. (2021). Confirmatory composite analysis using partial least squares: Setting the record straight. *Review of Managerial Science, 15*(5), 1311–1345. https://doi.org/10.1007/s11846-020-00405-0
Schuberth, F. (2023). The Henseler–Ogasawara specification of composites in structural equation modeling: A tutorial. *Psychological Methods, 28*(4), 843–859. https://doi.org/10.1037/met0000432
Schuberth, F., Hubona, G., Roemer, E., Zaza, S., Schamberger, T., Chuah, F., … Henseler, J. (2023a). The choice of structural equation modeling technique matters: A commentary on Dash and Paul (2021). *Technological Forecasting and Social Change, 194*, Article 122665. https://doi.org/10.1016/j.techfore.2023.122665
Schuberth, F., Rosseel, Y., Rönkkö, M., Trinchera, L., Kline, R. B., & Henseler, J. (2023b). Structural parameters under partial least squares and covariance-based structural equation modeling: A comment on Yuan and Deng (2021). *Structural Equation Modeling: A Multidisciplinary Journal, 30*(3), 339–345. https://doi.org/10.1080/10705511.2022.2134140
Schuberth, F., Schamberger, T., & Henseler, J. (2024). More powerful parameter tests? No, rather biased parameter estimates. Some reflections on path analysis with weighted composites. *Behavior Research Methods, 56*, 4205–4215. https://doi.org/10.3758/s13428-023-02256-5
Schuberth, F., Zaza, S., & Henseler, J. (2023d). Partial least squares is an estimator for structural equation models: A comment on Evermann and Rönkkö (2021). *Communications of the Association for Information Systems, 52*(1), 711–714. https://doi.org/10.17705/1CAIS.05232
Tenenhaus, M., Vinzi, V. E., Chatelin, Y.-M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics & Data Analysis, 48*(1), 159–205. https://doi.org/10.1016/j.csda.2004.03.005
Werts, C. E., Linn, R. L., & Jöreskog, K.G.. (1974). Intraclass reliability estimates: Testing structural assumptions. *Educational and Psychological Measurement, 34*(1), 25–33. https://doi.org/10.1177/001316447403400104
Yu, X., Schuberth, F., & Henseler, J. (2023). Specifying composites in structural equation modeling: A refinement of the Henseler–Ogasawara specification. *Statistical Analysis and Data Mining, 16*(4), 348–357. https://doi.org/10.1002/sam.11608
Yu, X., Zaza, S., Schuberth, F., & Henseler, J. (2021). Counterpoint: Representing forged concepts as emergent variables using composite-based structural equation modeling. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems, 52*(SI), 114–130. https://doi.org/10.1145/3505639.3505647