Data in Brief 48 (2023) 109074

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

A perspective on using partial least squares structural equation modelling in data articles☆

Christian M. Ringle[a], Marko Sarstedt[b,c], Noemi Sinkovics[d,*], Rudolf R. Sinkovics[e]

[a] *Hamburg University of Technology, Department of Management Sciences and Technology, Hamburg, Germany*
[b] *Ludwig-Maximilians-University Munich, Germany*
[c] *Babeş-Bolyai University, Romania*
[d] *University of Glasgow, Adam Smith Business School, United Kingdom*
[e] *University of Glasgow, Adam Smith Business School, United Kingdom, and LUT University, Lappeenranta, Finland*

ARTICLE INFO

Article history:
Received 9 November 2022
Revised 7 March 2023
Accepted 10 March 2023
Available online 21 March 2023

Dataset link: Empty reference data (filler link) (Original data)

Keywords:
PLS-SEM
Partial least squares
Structural equation modelling
Open science
Discriminant validity
HTMT

ABSTRACT

This perspective article on using partial least squares structural equation modelling (PLS-SEM) is intended as a guide for authors who wish to publish datasets that can be analysed with this method as stand-alone data articles. Stand-alone data articles are different from supporting data articles in that they are not linked to a full research article published in another journal. Nevertheless, authors of stand-alone data articles will be required to clearly demonstrate and justify the usefulness of their dataset. This perspective article offers actionable recommendations regarding the conceptualisation phase, the types of data suitable for PLS-SEM and quality criteria to report, which are generally applicable to studies using PLS-SEM. We also present adjusted versions of the HTMT metric for discriminant validity testing that broaden its applicability. Further, we highlight the benefit of linking data articles to already published research papers that employ the PLS-SEM method.

☆ Noemi Sinkovics has a role as Editor In Chief of this journal but had no involvement in the peer-review of this article and has no access to information regarding its peer-review.

* Corresponding author

E-mail addresses: c.ringle@tuhh.de (C.M. Ringle), sarstedt@lmu.de (M. Sarstedt), noemi.sinkovics@glasgow.ac.uk (N. Sinkovics), rudolf.sinkovics@glasgow.ac.uk (R.R. Sinkovics).

https://doi.org/10.1016/j.dib.2023.109074

1. Background

This perspective article on using partial least squares (PLS) for estimating relationships among observed and latent variables [1,2] – also referred to as PLS path modelling (e.g., [3,4]), the PLS approach to structural equation modelling [e.g., [5]], or PLS structural equation modelling (PLS-SEM; e.g., [6,7]) – in data articles has at least two objectives. *First*, it is intended as a guide for authors who wish to publish datasets that can be analysed with PLS-SEM as a stand-alone data article. Stand-alone data articles are not linked to an original research article published in another journal. These datasets are frequently generated through some form of replication of previous studies and thus may not meet the criterion of novel theoretical contribution as defined by many traditional journals in the broader field of business and management. However, it is important to emphasise that empirical contributions are no less important for the advancement of science than theoretical contributions [8]. Therefore, a stand-alone data article can be seen as valuable as long as it makes a meaningful empirical contribution. Köhler and Cortina [8] offer a typology of the various forms of replication and examine the contribution of each type. They distinguish between literal replication, quasi random replication, constructive replication, confounded replication, and regressive replication. The degree of the empirical contribution varies across the different forms. Specifically, the value of confounded replications and regressive replications are unclear as they contain shortcomings vis-à-vis the original article [8].

Second, this perspective article aims to highlight the benefits of publishing data articles alongside original research articles in other journals. Linked data articles can be regarded as an important enabler on the road towards 'gold standard reproducibility' [9]. Research articles that only describe the data but do not make the data publicly available are only partially reproducible and may not be fully replicable. This is partly due to the limited space available in research articles that can be dedicated to the methods section. Some journals have strict word count limits. Although it is possible to host supplementary files with research articles, the reader may not be able to easily understand them without additional explanation. Therefore, reproducibility requires publication of the research article that provides the theoretical underpinning and interpretation of the results, a detailed description of the method and research design, as well as the publication of the data set [9].

A linked data article allows authors to focus more on the theoretical framing and interpretation of results in the main research article and to offer a more detailed description of all the decisions that have been made in the research design and data analysis aspect of the process in the data article. Some decisions that hold relevance for future replication studies may not be reported in the original research article because of space constraints. Further, in linked data articles, authors can provide details of all the items in their dataset beyond the items analysed in the main research article.

Both types of data articles support the advent of open science. Open science is fundamentally changing the landscape in which research, education and innovation are undertaken, archived, curated and disseminated around the world [10]. Open science is about ensuring research findings are accessible to all rather than keeping them behind a paywall. It encompasses four principles or pillars of openness, namely, open data, open code, open papers and open reviews [9,11]. The case for open science has become a particularly strong one as the world grapples with global challenges of resource scarcity, climate change, poverty, ill-health, pollution, rapid urbanisation and food insecurity [12]. However, although open science practices are increasingly permeating the natural and medical sciences, other fields, including the social sciences, business and management, and economics, are lagging behind in adopting these practices [13–15]. As studies in business and management increasingly use the PLS-SEM method (e.g., [16]), this perspective article also aims to encourage researchers to keep abreast of reporting standards and data curating practices in other disciplinary areas. We therefore revisit PLS-SEM's key characteristics and reporting standards, pointing to contradicting streams in the methodological literature that authors need to be aware of. In our discussion of model evaluation metrics, we also offer adjusted versions of the HTMT metric for discriminant validity testing that broaden its applicability.

2. Recommendations

2.1. The conceptualisation phase: ensuring the model is meaningful and highlights interesting relationships

A key characteristic of appropriately anchoring any quantitative research paper is the alignment of concepts with the operational part of a study [17]. While data articles do not feature extensive theoretical or conceptual sections, it is nevertheless critically important to demonstrate the underpinning conceptual or theoretical 'story' explored in the data and its origins. This step includes a justification why specific concepts were selected and how they are related. The model specification should be described or represented by a graphical model illustration. In addition to journal articles, construct definitions and their measurements (i.e., sets of items or 'scales') can also be found in scale handbooks. Scale handbooks are available for sub-fields and functional areas of business and management, such as marketing [18,19] entrepreneurship [20], work organisation [21], and organisational behaviour or human resource management [22]. In addition, journal articles routinely document item wordings in the main text or an appendix. In reviewing existing scales, researchers should closely consider their psychometric properties with regard to prior reliability and validity assessments [23]. Regarding the underpinning of the measurement scales' quality, it is usually a signal of good quality if the constructs on which PLS path models are built originate from high level and reputed journal publications rather than outlets that are not widely known. Journal lists [24–26] offer guidance regarding the reputation and quality levels of published outputs, which improves the 'face validity' of the scales and items reported.

Advice to authors: Describe the concepts and their relationships and report the measurement scales, including underpinning literature. If the scales have been adapted, report the changes made to the scales (highlight the changes to the wording).

If the authors' path model has already been published, they will still need to demonstrate a high standard of care with the reporting and referencing of source materials. In addition, authors should strengthen the trustworthiness and cogency of the data that they seek to make publicly available. Authors will also need to explain how the data article adds value to the information provided in the original research article. This is important because a linked data article should not be a simple replication of the methods section of the original article.

In stand-alone submissions to a data journal, where conceptual and theoretical reference work that develops the reported PLS path model does not yet exist, authors need to build their case carefully to demonstrate their model's usefulness. It is not sufficient to show that the data used to deploy the PLS-SEM method satisfies the common model evaluation criteria as documented in, for example, Hair et al. (2022). Careful, transparent, and practically engaging writing is required to provide a convincing chain of arguments demonstrating why the model offers meaning. Since space is limited, this may involve references to a conceptual debate elsewhere that could be enriched through the dataset. A multi-stage pattern-matching process can be a useful way to ascertain that the conceptual model is meaningful and relevant in a specific context [27].

Our advice to authors: Demonstrate that the model is meaningful and relevant.

2.2. Model estimation: Why PLS-SEM?

As much as the conceptual or theoretical level and measurement level cannot be divorced [28], the write-up of a conceptualisation stage in a data article and the model estimation stage are intertwined and cannot be separated [29]. Therefore, a clear statement of the objectives of the analysis is required in addition to an explanation of how the PLS-SEM results support the accomplishment of the stated objectives. With respect to the associated empirical aspects, this relates to, for example, high model complexity, the estimation of formative measurement models, and the use of small samples (e.g., [7]). Most notably, researchers have emphasized that PLS-SEM is particularly well suited for estimating models from an explanation-prediction perspective (e.g., [30,31,32]), which implies an understanding of the relationships assumed in the model as well as its ability to predict theoretical concepts under consideration [33].

Our advice to authors: Discuss how the use of PLS-SEM contributes to the overall aim of the analysis.

Researchers have developed variants of the standard PLS-SEM algorithm to deal with different data or model types. For instance, to adjust the original parameter estimates to accommodate common factor models, authors may revert to consistent PLS (PLSc), "which was first mathematically developed by Theo K. Dijkstra" [114] see also p. 299, for instance, [34,35], its PLSe1/PLS2e extensions [36,37], or Yuan et al.'s [38] Cronbach α based approach. Compared to the standard algorithm, PLSc builds on different assumptions regarding the nature of measurement, most notably, how theoretical concepts are represented in statistical models [29]. While different perspectives are tenable [39], authors need to be aware of the underlying assumptions and potential consequences if there is a misalignment between the model type (i.e., composite vs. common factor models) and estimator (e.g., standard PLS-SEM vs. PLSc-SEM) [40,41].

Our advice to authors: Consider the assumptions, characteristics, and different outcomes of a composite estimation of constructs compared to a common factor estimation.

2.3. Types of data suitable for PLS-SEM and quality criteria to report

The PLS-SEM algorithm relies on a series of linear regressions coupled with linear combinations to estimate the model parameters [2], Chapter 2 [42]. The indicator variables in a construct's measurement model should have data on a metric scale (ratio or interval measurement), such as age or income [43], Chapter 1. However, researchers also use survey data with ordinal scales (e.g., Likert-type scales). These data are useful for PLS-SEM when the researchers can justify equidistant data points (i.e., measurement on a quasi-metric scale). Alternatively, researchers may consider approaches for nonmetric partial least squares [44] and ordinal partial least squares (e.g., [45,46]). Also, Lohmöller [2], Chapter 4) presents a PLS-SEM approach that only uses binary variables (see also [47]), which can also be used for categorical and ordinal variables, or a mixture of both. However, binary, and categorical variables should usually not be mixed with metric or quasi-metric indictors in a construct's measurement model as this may complicate the interpretation of the results. PLS-SEM users can relatively easily include binary-coded data as control variables, moderators or grouping variables (i.e., when conducting a PLS-SEM multigroup analysis) in their analyses [for further details see [48]].

Our advice to authors: Report the scale of the data.

The data used in the analysis can come in the form of primary or secondary data (for further details on the suitability and use of secondary data in PLS-SEM, see [43], Chapter 1). In using such data, researchers typically draw on samples, drawn from a larger population. This step, however, requires (1) clearly defining the population about which inferences will be made, and (2) statements regarding the sample's (specific) representativeness [49]. The latter step entails safeguarding that the sample matches the population with regard to relevant characteristics (e.g., age, income, or gender). Researchers sometimes use samples generated by market research companies (e.g., [50]). These data usually include a weighting variable to ensure the sample's representativeness with regard to relevant characteristics. The weighted PLS-SEM algorithm [51] – see also [52] – allows aligning the sample and population with regard to such characteristics by weighting observations differently. Further, the use of digital sources of data such as Amazon's Mechanical Turk (MTurk) is on the rise. Authors are advised to familiarize themselves with the discussions around the pitfalls of such data sources [53].

Our advice to authors: Discuss the population and the sample's structure. Apply sampling weights in case of a mismatch between sample and population with regard to key characteristics.

As with any multivariate analysis method, the use of large samples is usually advantageous. However, in some situations, the population of interest is extremely small. For example, in business-to-business and family business research, populations are often restricted in size (e.g., [54,55]). In such situations, researchers must ensure that the PLS-SEM method meets the minimum sample size requirements (e.g., [5,43], Chapter 1). Researchers using PLS-SEM should therefore run power analyses or rely on heuristics such as Kock and Hadaya [56] inverse square root method to determine the sample size needed for achieving a certain level of statistical power (typically 80%; [57]). Researchers in epidemiology have proposed alternative means for minimum sample size requirements that do not rely on the concept of statistical power, but specify maximum allowable margins of error in confidence intervals for population effect sizes based on sample data [58]. Such an approach may be advantageous as there is no reliance on traditional inference testing, which has come under scrutiny in some research areas (e.g., [59]).

Our advice for authors: Ensure that the analysis yields sufficient levels of statistical power.

Datasets typically have missing values, which need to be treated. As missing value treatment is, in a sense, data manipulation, it should remain within reasonable limits. For instance, researchers have suggested that variables with more than 15% missing values be removed from the dataset [43], Chapter 1. Typical missing value treatment options are mean replacement, expectation-maximisation (EM), hot and cold deck, maximum likelihood, nearest neighbour, and regression imputation [60], Chapter 2. For example, Wang, Lu and Liu [61] suggested an EM algorithm for missing data imputation in PLS-SEM. Alternatively, researchers can delete observations with missing values, a process also known as casewise or listwise deletion. For studies using PLS-SEM, Hair, Hult, Ringle and Sarstedt [43], Chapter 1 suggest applying mean replacement when less than 5% of the values per variable are missing and to consider the use of casewise deletion otherwise. However, researchers need to ensure that the deletion of observations does not occur systematically and that sufficient observations remain for analysis.

Our advice for authors: Report the percentage of missing values for the entire dataset and the percentage of missing values per variable. Additionally, report the missing value treatment option used.

Besides treating missing values, authors should inspect their dataset for suspicious response patterns, such as straight lining or alternating extreme pole responses. An analysis of descriptive statistics (e.g., mean, variance and distribution of the answers per respondent) and/or graphical representation of the data enables the identification of suspicious response patterns (for further details, see [43], Chapter 1). Researchers should also rule out possibility that participants respond at random to survey items or tasks. Prior research has shown that rates of random responding can be substantial, offering practical suggestions for detecting corresponding response patterns [113].

Our advice to authors: Report the analyses carried out to identify and treat suspicious response patterns, including random responding.

Similarly, researchers typically address outliers, which can come in the form of extreme responses or extreme values. Statistical software packages support the identification of outliers through univariate, bivariate or multivariate statistics and charts, for example, box plots and stem-and-leaf plots [62], Chapter 5. Boxplots can help identify responses that are three times the interquartile range below the first quartile or above the third quartile. Such responses could be extreme outliers due to erroneous data in which case they need to be eliminated. If an explanation is available for exceptionally high or low values, outliers are typically retained; see Sarstedt and Mooi [62], Chapter 5, for more details about outliers and their treatment.

Our advice for authors: Report the outlier detection approach, provide a justification for deleting outliers, and report the number of outliers deleted.

Finally, researchers should provide the covariance matrix of their data as well as key descriptive statistics. The latter should include, for example, the minimum and maximum value, the mean value and median, the skewness and kurtosis, and the result of a nonnormality test statistic (e.g., the Anderson-Darling test or Cramér-von Mises test). PLS-SEM is a nonparametric method, which can accommodate nonnormal data. However, nonnormal data may lead to skewed bootstrap distributions, which negatively affect the bootstrap confidence intervals for significance testing. In such a situation, researchers may prefer bias-corrected and accelerated bootstrap confidence intervals over percentile-based confidence intervals – the standard variant in PLS-SEM studies [63].

Our advice for authors: Provide the covariance matrix and key descriptive statistics, including results of nonnormality tests.

2.4. Everything PLS-SEM, once the model stands

Model estimation using PLS-SEM requires reporting and citing the software used as well as any algorithm settings deviating from the defaults; not only for the PLS-SEM algorithm, but all computations including, for instance, bootstrapping and $\text{PLS}_{\text{predict}}$. Besides commercial software packages such as SmartPLS [64] and WarpPLS [65], authors can also revert to the open source statistical software R [66] and its cSEM [67] and SEMinR [68] packages for PLS-SEM.

After the results computation, authors should report the criteria for reflective and formative measurement model assessment and the structural model. Several articles and textbooks provide overviews of which PLS-SEM results and criteria should be reported and how [69–71]. Table 1 provides a summary of key criteria as discussed in, for example, Hair, Hult, Ringle and Sarstedt [43]. Additional components may include endogeneity assessment (e.g., [72]), model comparisons (e.g., [31]), or extended visualizations of PLS-SEM results such as importance-performance map analysis (e.g., [73]).

In terms of measurement model assessment, particular care should be devoted to discriminant validity assessment, which ensures that each construct is empirically unique and captures a phenomenon not represented by other constructs in a statistical model [74]. Safeguarding discriminant validity is of crucial concern to ensure that implications drawn from the analysis of structural model relationships are not the result of statistical discrepancies [75]. The primary criterion for PLS-SEM-based discriminant validity assessment is Henseler et al.'s [76] HTMT criterion and its recent extension HTMT2 [77]. While both criteria will not differ much in common applications, their results may be affected by negative correlation patterns. We therefore propose adjusted versions of the criteria, which rely on absolute correlations in their computations. We refer to these criteria as HTMT+ and HTMT2+ to indicate that they only employ positive (absolute) correlation values (hence, the additional plus symbol). We document these criteria and showcase their usefulness in the Appendix. To assess discriminant validity, researchers are advised to use the adjustment of the HTMT criterion (i.e., HTMT+), which is available, for instance, in the widely adopted SmartPLS software [64] for PLS-SEM analyses. In some extreme data and model constellations (especially when the number of indicators is low and their heterogeneity of loadings is particularly high), researchers may also consider the adjustment of the HTMT2 criterion (i.e., HTMT2+).

In terms of structural model assessment, authors need to assess their model's predictive power, which is often neglected in regression-based studies [32]. Authors are therefore advised to routinely run Shmueli et al.'s [78] $\text{PLS}_{\text{predict}}$ procedure and follow extant guidelines for results reporting [79]. While the $\text{PLS}_{\text{predict}}$ focuses on the model's predictive power on an item-level, authors are also advised to compare their model's predictive power on a construct level – either on the grounds of all endogenous constructs or one specific key target construct. Sharma et al.'s

Table 1
Results reporting.

Evaluation purpose	Criterion
Reflective measurement models	
Internal consistency reliability	$\rho_A \geq 0.7$ and < 0.95; consider reporting Cronbach's α and ρ_C as lower and upper boundaries, respectively (note that these two additional metrics require that indicator correlations in a measurement model are either all positive or negative)
Indicator reliability	Loadings: ≥ 0.7 (or 0.708 to be precise)
Convergent validity	Average variance extracted (AVE): ≥ 0.5
Discriminant validity	HTMT+ (and HTMT2+): ≤ 0.85 or 0.9. Assess whether the selected threshold falls into the (one-sided) bootstrap confidence interval
Formative measurement models	
Convergent validity	Redundancy analysis
Collinearity	Variance inflation factor (VIF): ≤ 5, ideally ≤ 3
Significance and relevance of indicators	Bootstrap-based significance testing of indicator weights (and loadings); size of the coefficients
Structural model	
Collinearity	Variance inflation factor (VIF): ≤ 5, ideally ≤ 3
Significance and relevance of path coefficients	Bootstrap-based significance testing; (effect) size of the coefficients
Predictive power and model fit	Primary focus is on *prediction*: Focus on predictive power assessment using $\text{PLS}_{\text{predict}}$ and the cross-validated predictive ability test (CVPAT) Primary focus is on *explanation*: Model fit metrics (e.g., SRMR) and bootstrap-based tests for model fit, but consider limitations related to their applicability Focus is on *both* prediction and explanation: Consider the trade-off between model fit and predictive power

[80] extension of Liengaard et al.'s [31] cross-validated predictive ability test (CVPAT) facilitates such analyses.

In following the guidelines in Table 1, authors need to be aware of two streams in the PLS-SEM literature that suggest different procedures for model evaluation. While Schuberth, Rademaker and Henseler [81] emphasize the role of model fit for results assessment, using indices such as the GFI, NFI, and SRMR or bootstrap-based tests for model fit, Hair, Hult, Ringle and Sarstedt [43], Chapters 1 and 6 focus on predictive power assessment, emphasizing the method's causal-predictive nature (see also [82]). Both procedures are tenable, but authors need to be aware of the two streams and make an informed assessment which route they follow. In addition, researchers need to be aware of potential trade-offs between the two perspectives. Focusing on model fit may sacrifice predictive power [32] – a practice which has been intensively discussed in different methodological contexts [83,84]. Finally, researchers can conduct various robustness checks to safeguard the validity of the results. For instance, Hair et al. [85], Sarstedt et al. [69], and Sarstedt et al. [86] provide an overview of these analyses and the reporting of their results.

Our advice for authors: Comply with the most recent guidelines for PLS-SEM use, while considering different streams in the literature. Run robustness checks to safeguard the validity of the results.

2.5. How to communicate the value to user communities

Authors of data articles are expected to articulate why and how the data article generates value. [8]. Some generic examples to aid the communication of the value statement are provided here:

Reference material: A dataset can be useful as a reference for understanding a particular phenomenon in a particular context and can illuminate the hands-on application of specific methodologies that were used to uncover the phenomenon. This can be of particular relevance for single-shot (cross-sectional) primary and survey-based data, which are usually not available via public databases or research archives of funding bodies.

Replication: If the underlying dataset was produced through replication, authors need to be aware which form of replication they applied [8]. This will aid the communication of the dataset's contribution. With the provision of openly accessible data, it is possible to establish whether previously published empirical findings can be repeated and whether similar findings can be obtained under comparable conditions. When using the same data, this refers to the stability of results through verification or reanalysis and thus points to the findings' robustness. As soon as additional and different data are introduced, further value can be provisioned through direct replication and extension [14], Table 9.1 on p. 159. Such replications are crucial in light of increasing concerns regarding the stability of even seminal behavioural effects like nudging [87]. Authors have the opportunity to use the data article to highlight the research avenues that the dataset feeds into as well as delineate other potential uses for the dataset. Some of this may stem from 'future research' sections of previous research articles or relate to new possibilities that the new data provide.

Data pooling: Pooling data from previously existing studies can be useful to increase the sample size and the statistical power of the analysis. In international contexts, the ability to pool data can support the development of cross-country, comparative models and thus help establish an understanding of key differences. Conceptual issues of equivalence must be considered, which include research subject, time and context, to render such data pooling efforts useful [88]. Empirical considerations and words of caution regarding equivalence issues should be discussed in the model evaluation section. In the context of PLS-SEM, this involves running Henseler et al.'s [89] MICOM procedure.

Temporal value: The provision of open access data may contribute to an understanding of value changes regarding the underpinning constructs over time. This will be of special value to studies in which sociological or psychological constructs are involved. While different sampling and equivalence concerns may render direct comparison difficult, tendential assessment can produce value for research communities.

Frequently, authors become overly excited when referring to their 'unique' datasets, even when these datasets would not have been possible without the help of intermediary parties and grant-funding bodies and may have been the result of significant collaborative efforts involving academics at various stages of their careers. Researchers in the social sciences, humanities, and business and management are still considerably reluctant to share their data [90]. Christensen, Freese and Miguel [14] report the case of a mysterious psychologist who claimed a loss of data in response to a query about suspicious results, which "exemplifies one reason researchers might not want to share their data: *maybe they have something to hide*" [14], p. 176. Null results in relationships not openly documented in the original manuscript may be such a reason. In support of open science, *data journals,* encourage the inclusion of null results in data articles. Another, less uncomfortable explanation for the reluctance of researchers to share datasets is the fear of losing further publication opportunities to extend the model and analysis, if other research teams manage to exploit these opportunities more quickly. Yet, with reproducibility and transparency requests increasing [91] and social science, humanities and business and management becoming eager to close the open data gap to the natural and medical sciences [92,93], it will also become attractive for curators of such datasets to receive appropriate recognition and credit. New citation standards regarding referencing of datasets, such as those included in the seventh edition of the *Publication Manual of the American Psychological Association* [94] and modern 'how to' guides that include datasets and data articles [95], will help increase the value of open data further.

3. Key points

- Ensure model is meaningful: Report the model, measures, including any changes made and underpinning literature
- Demonstrate that the model and PLS-SEM results are meaningful and relevant
- Check whether the data is appropriate for analysis and report quality criteria
- Report scale levels of data, representativeness of sample, minimum sample size requirements, percentage of missing values and missing data option used, analyses carried out to identify suspicious response patters and deleted observations, covariance matrix and descriptive statistics, results reporting including criteria for reflective and formative measurement models and the structural model.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Although this research does not use the statistical software SmartPLS (https://www.smartpls.com), Christian M. Ringle acknowledges a financial interest in SmartPLS.

Data Availability

Empty reference data (filler link) (Original data) (https://www.sciencedirect.com/journal/data-in-brief).

A. Appendix

A.1. Extending the standard HTMT metrics for discriminant validity testing

For any two constructs ξ_i and ξ_j measured by K_i and K_j indicators, Henseler, Ringle and Sarstedt [76] defined the HTMT as the arithmetic mean of the heterotrait-heteromethod correlations $r_{ig,jh}$ relative to the geometric mean of the average monotrait-heteromethod correlations of ξ_i (i.e., $r_{ig,ih}$) and the average monotrait-heteromethod correlation of ξ_j (i.e., $r_{jg,jh}$):

$$\mathrm{HTMT}_{ij} = \frac{\frac{1}{K_i K_j}\sum_{g=1}^{K_i}\sum_{h=1}^{K_j} r_{ig,jh}}{\sqrt{\frac{2}{K_i(K_i-1)}\sum_{g=1}^{K_i-1}\sum_{h=g+1}^{K_i} r_{ig,ih}\cdot\frac{2}{K_j(K_j-1)}\sum_{g=1}^{K_j-1}\sum_{h=g+1}^{K_j} r_{jg,jh}}}.$$

Based on prior research and simulation study results, Henseler, Ringle and Sarstedt [76] suggest a threshold value of 0.90 if constructs are conceptually very similar and 0.85 if the constructs are conceptually more distinct. That is, an HTMT value for two constructs that exceeds the corresponding threshold indicates a lack of discriminant validity. In addition, these authors suggested testing whether HTMT is significantly different from unity, which is a very liberal way of testing for discriminant validity testing. Rather than relying on cutoff values lower than unity, Franke and Sarstedt [74] suggest testing whether an HTMT value differs significantly from a selected threshold different from unity (e.g., 0.90). To do so, researchers can compute the HTMT statistic's bootstrap-based confidence interval and assess whether the corresponding cutoff value falls into the interval.

The HTMT criterion rests on the assumption that the indicators' population loadings are equal, a requirement also referred to as tau-equivalence [96]. However, this assumption is unlikely to hold in many empirical settings [97]. Addressing this concern, Roemer, Schuberth and

Table A1
Indicator correlations (wave #1).

		ξ_1				ξ_2			
		x_{11}	x_{12}	x_{13}	x_{14}	x_{21}	x_{22}	x_{23}	x_{24}
ξ_1	x_{11}	1.000							
	x_{12}	0.504	1.000						
	x_{13}	0.456	0.632	1.000					
	x_{14}	0.463	0.485	0.510	1.000				
ξ_2	x_{21}	0.311	0.477	0.380	0.289	1.000			
	x_{22}	0.396	0.468	0.405	0.591	0.380	1.000		
	x_{23}	0.362	0.435	0.382	0.341	0.445	0.396	1.000	
	x_{24}	0.343	0.455	0.367	0.314	0.556	0.400	0.382	1.000

Note: The dark-shaded elements denote the heterotrait-heteromethod correlations, while the remaining off-diagonal elements correspond to the monotrait-heteromethod correlations.

Henseler [77] introduced an extension of the original criterion, which considers the geometric mean instead of the arithmetic mean for calculating the average indicator correlations [77], Eq. 3[1]:

$$\mathrm{HTMT2}_{ij} = \frac{\sqrt[K_i \cdot K_j]{\prod_{g=1}^{K_i} \prod_{h=1}^{K_j} r_{ig,jh}}}{\sqrt{\sqrt[\frac{K_i^2 - K_i}{2}]{\prod_{g=1}^{K_i - 1} \prod_{h=g+1}^{K_i} r_{ig,ih}} \cdot \sqrt[\frac{K_j^2 - K_j}{2}]{\prod_{g=1}^{K_j - 1} \prod_{h=g+1}^{K_j} r_{jg,jh}}}}$$

Roemer et al.'s [77] simulation study shows that the HTMT2 slightly outperforms the original metric when the indicator loading patterns are extremely heterogeneous (e.g., some indicator loadings are around 0.5, while others are close to 1), particularly when ξ_i and ξ_j are highly correlated (see the HTMT2 discussion in [69]). While the HTMT2 addresses conceptual concerns regarding the original metric's tau-equivalence assumption, both HTMT variants require the average monotrait-heteromethod correlations to be positive: "Computation of the HTMT/HTMT2 assumes that all intra-block and inter-block correlations between indicators are either all-positive or all-negative" [67], p. 17. Similarly, Roemer et al. [77], p. 2647 note that the HTMT "is not defined for cases, in which indicator correlations are negative." However, this is not necessarily the case in empirical applications of PLS-SEM and SEM in general.

Consider the following example taken from a model on customer satisfaction, where the construct ξ_1, represents the respondents' trust [98] and is measured with indicators x_{11}, x_{12}, x_{13}, and x_{14}. The construct ξ_2 represents their perceived switching costs [99] and is measured with indicators x_{21}, x_{22}, x_{23}, and x_{24}. A professional market research firm collected responses from 5,368 different consumers in three parallel study waves. Each study wave drew random samples from the German population, but with slight variations in the indicator scaling. Specifically, construct measurement in wave #1 (n=1,812) relies on the original indicator wording; in wave #2 (n=1,767), x_{11} was reverse-scaled; in wave #3 (n=1,789), x_{11} and x_{21} were reverse-scaled.

Table A1 documents the correlation patterns between the indicators of ξ_1 and ξ_2 for wave #1. The dark-shaded elements denote the heterotrait-heteromethod correlations, while the remaining off-diagonal elements correspond to the monotrait-heteromethod correlations of ξ_1 and ξ_2 respectively.

All indicator correlations are positive, producing the following values for HTMT and HTMT2:

$$\mathrm{HTMT} = \frac{0.395}{\sqrt{0.508 \cdot 0.427}} = 0.848$$

[1] Note that the denominator in Roemer et al.'s [77] Eq. 3 erroneously denoted the indicator correlations of ξ_i as $r_{jg,jh}$ and the upper index of the corresponding product term (i.e., in the second position) with K_j. We corrected these errors in our definition of $\mathrm{HTMT2}_{ij}$.

Table A2
Indicator correlations (wave #2).

		ξ_1				ξ_2			
		x_{11}	x_{12}	x_{13}	x_{14}	x_{21}	x_{22}	x_{23}	x_{24}
ξ_1	x_{11}	1.000							
	x_{12}	-0.520	1.000						
	x_{13}	-0.446	0.659	1.000					
	x_{14}	-0.426	0.499	0.501	1.000				
ξ_2	x_{21}	-0.317	0.526	0.420	0.317	1.000			
	x_{22}	-0.413	0.506	0.419	0.621	0.397	1.000		
	x_{23}	-0.369	0.442	0.358	0.339	0.438	0.419	1.000	
	x_{24}	-0.362	0.549	0.441	0.388	0.579	0.460	0.419	1.000

Note: The dark-shaded elements denote the heterotrait-heteromethod correlations, while the remaining off-diagonal elements correspond to the monotrait-heteromethod correlations.

Table A3
Indicator correlations (wave #3).

		ξ_1				ξ_2			
		x_{11}	x_{12}	x_{13}	x_{14}	x_{21}	x_{22}	x_{23}	x_{24}
ξ_1	x_{11}	1.000							
	x_{12}	-0.449	1.000						
	x_{13}	-0.373	0.615	1.000					
	x_{14}	-0.404	0.487	0.510	1.000				
ξ_2	x_{21}	0.254	-0.469	-0.375	-0.309	1.000			
	x_{22}	-0.343	0.437	0.382	0.627	-0.350	1.000		
	x_{23}	-0.276	0.422	0.331	0.343	-0.448	0.373	1.000	
	x_{24}	-0.350	0.488	0.390	0.355	-0.546	0.384	0.407	1.000

Note: The dark-shaded elements denote the heterotrait-heteromethod correlations, while the remaining off-diagonal elements correspond to the monotrait-heteromethod correlations.

$$HTMT2 = \frac{0.388}{\sqrt{0.505 \cdot 0.423}} = 0.840$$

Assuming a cutoff value of 0.90 or higher in light of the constructs' conceptual similarity, both HTMT metrics would indicate discriminant validity as both values are lower than this threshold. As expected in light of the measurement models' quality, HTMT and HTMT2 are in close correspondence [69]. Researchers could also run bootstrapping and construct confidence intervals to test whether the two HTMT values are significantly lower than this threshold.

However, a different picture emerges when computing the HTMT metrics based on the wave#2 correlation matrix. Using the indicator correlations documented in Table A2 as input produces an HTMT value far above 1, which is surprising as the reverse-coding of a single indicator (x_{11}) should not alter whether or not ξ_1 and ξ_2 are empirically distinct. More severely, HTMT2 is not defined because of the negative first terms that result in the formulas' denominator.

$$HTMT = \frac{0.242}{\sqrt{0.045 \cdot 0.452}} = 1.703$$

$$HTMT2 = \frac{0.416}{\sqrt{\sqrt[6]{-0.016} \cdot \sqrt[6]{0.008}}} = not\ defined$$

A similar problem occurs when using the indicator correlations from wave #3, where one indicator in each of the two constructs is reverse-coded (x_{11} and x_{21}). Using the indicator correlations documented in Table A3 as input produces negative values in the formulas' denominators, rendering the computation of HTMT and HTMT2 impossible:

$$\mathrm{HTMT} = \frac{0.119}{\sqrt{0.064 \cdot (-0.030)}} = not\ defined$$

$$\mathrm{HTMT2} = \frac{0.375}{\sqrt{\sqrt[6]{-0.010} \cdot \sqrt[6]{-0.005}}} = not\ defined$$

These and similar problems (e.g., negative and positive correlations cancelling each other out in the computation of the mean) have been recognized in the context of the original HTMT metric and considered in its implementation in SmartPLS, the most frequently used software for PLS-SEM analyses in various fields of business research (e.g., [100,101]) and other disciplines (e.g., [102,103]). Specifically, since the HTMT computation update in version 3.2.1 of SmartPLS 3 [104][2] and continuing with SmartPLS 4 [64], the software considers the absolute correlation values in the computation of the HTMT.[3] Since the correlations are only intended to determine the indicators' empirical overlap, their signs are not decisive. Hence, both the HTMT and HTMT2 metrics may draw on absolute correlations to avoid problems that occur when encountering correlation patterns as in the wave #2 and wave #3 examples.

To distinguish these adjusted versions from their previous implementations, we refer to these metrics as HTMT+ and HTMT2+:

$$\mathrm{HTMT}+_{ij} = \frac{\frac{1}{K_i K_j} \sum_{g=1}^{K_i} \sum_{h=1}^{K_j} \left| r_{ig,jh} \right|}{\sqrt{\frac{2}{K_i(K_i-1)} \sum_{g=1}^{K_i-1} \sum_{h=g+1}^{K_i} \left| r_{ig,ih} \right| \cdot \frac{2}{K_j(K_j-1)} \sum_{g=1}^{K_j-1} \sum_{h=g+1}^{K_j} \left| r_{jg,jh} \right|}}$$

$$\mathrm{HTMT2}+_{ij} = \frac{\sqrt[K_i \cdot K_j]{\prod_{g=1}^{K_i} \prod_{h=1}^{K_j} \left| r_{ig,jh} \right|}}{\sqrt{\sqrt[\frac{K_i^2-K_i}{2}]{\prod_{g=1}^{K_i-1} \prod_{h=g+1}^{K_i} \left| r_{ig,ih} \right|} \cdot \sqrt[\frac{K_j^2-K_j}{2}]{\prod_{g=1}^{K_j-1} \prod_{h=g+1}^{K_j} \left| r_{jg,jh} \right|}}}$$

Since all correlations in wave #1 are positive, the results of HTMT+ and HTMT2+ are identical to those of HTMT and HTMT2. Computing the HTMT+ and HTMT2+ values for waves #2 and #3 as outlined above produces the following results, which are very close to the values from wave #1. This result is expected considering that the samples have been drawn from the same population.

Wave #2:

$$\mathrm{HTMT}+ = \frac{0.424}{\sqrt{0.509 \cdot 0.452}} = 0.885$$

$$\mathrm{HTMT2}+ = \frac{0.416}{\sqrt{0.503 \cdot 0.448}} = 0.876$$

Wave #3:

$$\mathrm{HTMT}+ = \frac{0.384}{\sqrt{0.473 \cdot 0.418}} = 0.865$$

$$\mathrm{HTMT2}+ = \frac{0.375}{\sqrt{0.467 \cdot 0.413}} = 0.854$$

This adjustment resolves the problem that a simple recoding of indicators or lack thereof may alter the results of a discriminant validity analysis.

[2] Version 3.2.1 of SmartPLS 3 was released on May 5, 2015 (see https://www.smartpls.com/release_notes/). Since then, the improved HTMT computation with absolute correlation values (i.e., HTMT+) has been documented on the SmartPLS webpages (https://www.smartpls.com/documentation/algorithms-and-techniques/discriminant-validity-assessment).

[3] The cSEM R package, which provides the HTMT and HTMT2 metrics, issues a warning when such inadmissible values occur [67].

In case of deliberately reverse-coded items, researchers may, of course recode the corresponding item(s) and apply the standard HTMT metrics [77]. However, negative correlation patterns may occur systematically, but without any prior expectations, causing the standard HTMT metrics to fail. For example, negative correlation patterns among indicators routinely occur in applications of personality scales and measurements of emotions (e.g., PANAS; [105]), clinical symptoms (e.g., [106,107]), response styles (e.g., [108,109]), and response times (e.g., [110])—to name a few. In such situations, reverse-coding might be difficult to defend from a conceptual standpoint, calling for metrics that are universally applicable to situations where negative correlation patterns occur. Note that the above examples are different from unsystematic variations in indicator loadings, for example, due to sampling variance. We will discuss this aspect in the following section.

A.2. Additional insights from simulations

Supplementary simulation studies will allow us to gain additional insights into how HTMT in its different variants can be used for discriminant validity testing. More specifically, one of the reviewers of this article pointed out that negative correlations may also occur unsystematically, despite positive population values due to sampling variation.[4] To show that such negative correlations can have adverse consequences for the computation of the HTMT+ metric, the reviewer presented the results of a small simulation study assuming a two-construct model with a 0.3 correlation, each measured by three indicators with loadings of 0.5. When estimating this model with $n = 100$ and 100,000 replications, the HTMT+ produced a larger than the expected value of 0.3, while the regular HTMT was not biased upwards.

While the metric's behaviour is important to understand, this simulation design constitutes a borderline situation with limited practical relevance due to (1) very low indicators loadings, (2) low construct correlations, and (3) a small sample size. We expect that the issue described above will vanish with higher indicator loadings, larger correlations between the constructs, and higher sample sizes. Taking these limitations into account, we first replicated the simulation with two constructs correlated at 0.3 and each measured by four indicators, which aligns with our empirical example presented in the previous section and corresponds to the measurement model complexity commonly encountered in applications of PLS-SEM (e.g., [69]). Our simulation study additionally considers the HTMT2 and HTMT2+ metrics as well as a series of further simulation factors. First, we considered a broader range of loadings. Specifically, loadings of 0.5 would raise a red flag in any empirical application (Table 1) as they would produce unbearably low internal consistency reliability (and convergent validity) values, rendering the discriminant validity assessment meaningless, because "a necessary (but not sufficient) condition for validity of measures is that they are reliable" [111], p. 6. We therefore also considered homogeneous loadings of 0.7 and a pattern of heterogeneous loadings with values of 0.9, 0.8, 0.7, and 0.6 for both constructs, ensuring that the lowest loading still meets minimum standards in terms of indicator reliability (Table 1). We reiterated these patterns by assuming a negative loading of equal size for one of the two constructs (e.g., -0.7 for the first indicator of the first constructs, all other loadings 0.7). Second, we not only considered a sample size of 100, but also 200 and 300, because prior reviews of PLS-SEM report similar average sample sizes in empirical applications of the method (e.g., [69]). We additionally considered a sample size of 10,000 to shed light on the metrics' asymptotic behaviour. Finally, as 10,000 replications produced highly similar results as 100,000 replications in our initial analysis, we considered this lower number to save compu-

[4] We would like to thank an anonymous reviewer for this helpful remark and the code provided for the statistical software R [66], which we modified and extended for our additional analyses.

Table A4
Simulation results (construct correlation 0.3, positive loadings only).

Sample size	HTMT	Number of inadmissible solutions	HTMT+	Number of inadmissible solutions	HTMT2	Number of inadmissible solutions	HTMT2+	Number of inadmissible solutions
Expected HTMT value: 0.3; all loadings 0.5								
100	0.305	0	0.420	0	0.329	5173	0.330	0
200	0.299	0	0.345	0	0.277	4715	0.267	0
300	0.301	0	0.323	0	0.271	4308	0.256	0
10,000	0.300	0	0.300	0	0.298	0	0.298	0
Expected HTMT value: 0.3; all loadings 0.7								
Sample size	HTMT	Number of inadmissible solutions	HTMT+	Number of inadmissible solutions	HTMT2	Number of inadmissible solutions	HTMT2+	Number of inadmissible solutions
100	0.300	0	0.313	0	0.295	2959	0.266	0
200	0.298	0	0.300	0	0.288	1341	0.273	0
300	0.299	0	0.300	0	0.289	519	0.282	0
10,000	0.300	0	0.300	0	0.299	0	0.299	0
Expected HTMT value: 0.3; loadings (per construct): 0.9 / 0.8 / 0.7 / 0.6								
Sample size	HTMT	Number of inadmissible solutions	HTMT+	Number of inadmissible solutions	HTMT2	Number of inadmissible solutions	HTMT2+	Number of inadmissible solutions
100	0.302	0	0.311	0	0.299	2629	0.273	0
200	0.300	0	0.302	0	0.292	1264	0.278	0
300	0.301	0	0.302	0	0.291	567	0.285	0
10,000	0.302	0	0.302	0	0.299	0	0.299	0

tational time.[5] We used the statistical software R [66] with its foreach package [112] for these simulations. Table A.4 shows the results of our analysis for measurement models with positive loadings only.

For loadings of 0.5, we find that the HTMT2 produces a great share of inadmissible solutions of up to 52% at sample sizes of 100, 200, and 300, while this is not the case for the other metrics. Looking at the estimates, we find that only HTMT returns (almost) unbiased average results for sample sizes of 100, 200, and 300. The other metrics improve in performance as sample sizes increases. Considering the scenario with 0.7 loadings for all indicators slightly changes the picture. HTMT2 still produces inadmissible solutions at small sample sizes, ranging between approximately 5% and 30%, which is less than in the previous scenario with lower loadings. In terms of parameter bias, HTMT shows practically no bias, whereas the HTMT+ metric is slightly upwards biased only for sample size of 100. HTMT2 and HTMT2+ are downwards biased at small sample sizes of 200 and 300, but less so compared to the scenario with 0.5 loadings. These metrics' bias diminishes as sample size increases. Finally, in the case of heterogeneous loadings, the HTMT2's share of inadmissible solutions is still high. The values of HTMT, HTMT+, and HTMT2 are highly similar, whereas HTMT2+ has a more pronounced underestimation tendency.

Table A.5 shows the results for the scenario where one loading is negative. As expected, HTMT and HTMT2 produce a significant share of inadmissible solutions. In the few cases where

[5] For instance, the first row in Table A.4 shows the outcomes for an expected average HTMT value of 0.3, all loadings set to 0.5 (i.e., four indicators per construct), 100 observations, and 10,000 replications. The corresponding results for 100,000 replications were (number of inadmissible solutions in squared brackets): HTMT = 0.305 [0], HTMT+ = 0.419 [0], HTMT2 = 0.329 [52,019], and HTMT2+ = 0.330 [0].

Table A5
Simulation results (construct correlation 0.3, negative and positive loadings).

Expected HTMT value: 0.3; first loading of the first construct -0.5, all other loadings 0.5									
Sample size	HTMT	Number of inadmissible solutions	HTMT+	Number of inadmissible solutions	HTMT2	Number of inadmissible solutions	HTMT2+	Number of inadmissible solutions	
100	0.689	5083	0.420	0	0.430	9854	0.330	0	
200	0.800	5015	0.345	0	0.326	9998	0.267	0	
300	0.893	5060	0.323	0	n/a	10,000	0.256	0	
10,000	2.252	4966	0.300	0	n/a	10,000	0.298	0	
Expected HTMT value: 0.3; first loading of the first construct -0.7, all other loadings 0.7									
Sample size	HTMT	Number of inadmissi-ble solutions	HTMT+	Number of inadmissible solutions	HTMT2	Number of inadmissible solutions	HTMT2+	Number of inadmissible solutions	
100	1.095	5095	0.313	0	n/a	10,000	0.266	0	
200	1.312	5079	0.300	0	n/a	10,000	0.273	0	
300	1.401	5070	0.300	0	n/a	10,000	0.282	0	
10,000	3.339	5105	0.300	0	n/a	10,000	0.299	0	
Expected HTMT value: 0.3; loadings: -0.9 / 0.8 / 0.7 / 0.6 (first construct), 0.9 / 0.8 / 0.7 / 0.6 (second construct)									
Sample size	HTMT	Number of inadmissi-ble solutions	HTMT+	Number of inadmissible solutions	HTMT2	Number of inadmissible solutions	HTMT2+	Number of inadmissible solutions	
100	n/a	10,000	0.311	0	n/a	10,000	0.273	0	
200	n/a	10,000	0.302	0	n/a	10,000	0.278	0	
300	n/a	10,000	0.302	0	n/a	10,000	0.285	0	
10,000	n/a	10,000	0.302	0	n/a	10,000	0.299	0	

these metrics can be computed, the results are substantially biased. On the contrary, the HTMT+ and HTMT2+ produce the same results as when assuming only positive loadings (Table A.4).

The previous settings assumed a relatively low construct correlation of 0.3. Understanding the metrics' performance in such a scenario is important as low construct correlations favour the occurrence of negative indicator correlations, triggering a substantial number of inadmissible solutions in the HTMT2 metric. At the same time, however, discriminant validity issues are less likely emerge when construct correlations are low. This is also why Roemer et al. [77] explicitly disregarded such correlations and have not identified HTMT2's tendency to produce inadmissible solutions. We therefore replicated the above analyses, but—in view of empirical applications—assumed a very high construct correlation of 0.8. Table A6 shows the results of our analysis for measurement models with positive loadings only.

As expected, we find that the number of inadmissible solutions in HTMT2 is much smaller than in the previous analyses, occurring mostly in situations where $n = 100$. Considering the parameter bias, we find that all metrics are generally in sync when the measurement models meet common quality standards. In these situations, HTMT2 has a marginal underestimation tendency, while HTMT has a marginal overestimation tendency. This also holds for the case of heterogeneous loadings, where one would expect the HTMT2 to outperform the HTMT metric. The HTMT+ and HTMT2+ metrics generally correspond to their standard versions, except for the borderline case with loadings of 0.5 and $n = 100$, where the differences are somewhat more pronounced.

A more detailed analysis of the results shows that HTMT and HTMT+ values are slightly too high (0.805 or 0.806 instead of 0.800) when loadings are heterogeneous, which parallels Roemer et al.'s [77] findings. However, in situation with equal loadings (in line with HTMT's assumption of tau-equivalent measurement models; [76,96]), this slight bias vanishes. On the contrary,

Table A6
Simulation results (construct correlation 0.8, positive loadings only).

Sample size	HTMT	Number of inadmissible solutions	HTMT+	Number of inadmissible solutions	HTMT2	Number of inadmissible solutions	HTMT2+	Number of inadmissible solutions
Expected HTMT value: 0.8; all loadings 0.5								
100	0.815	0	0.821	0	0.788	2457	0.774	0
200	0.804	0	0.804	0	0.783	362	0.778	0
300	0.804	0	0.804	0	0.789	34	0.788	0
10,000	0.800	0	0.800	0	0.800	0	0.800	0
Expected HTMT value: 0.8; all loadings 0.7								
Sample size	HTMT	Number of inadmissible solutions	HTMT+	Number of inadmissible solutions	HTMT2	Number of inadmissible solutions	HTMT2+	Number of inadmissible solutions
100	0.801	0	0.801	0	0.792	3	0.792	0
200	0.799	0	0.799	0	0.795	0	0.795	0
300	0.800	0	0.800	0	0.798	0	0.798	0
10,000	0.800	0	0.800	0	0.800	0	0.800	0
Expected HTMT value: 0.8; loadings (per construct): 0.9 / 0.8 / 0.7 / 0.6								
Sample size	HTMT	Number of inadmissible solutions	HTMT+	Number of inadmissible solutions	HTMT2	Number of inadmissible solutions	HTMT2+	Number of inadmissible solutions
100	0.806	0	0.806	0	0.793	28	0.793	0
200	0.805	0	0.805	0	0.795	0	0.795	0
300	0.806	0	0.806	0	0.798	0	0.798	0
10,000	0.806	0	0.806	0	0.800	0	0.800	0

HTMT2 and HTMT2+ are slightly too low at sample sizes of 100, 200, and 300, regardless of whether loadings are homogeneous or heterogeneous. This downward bias diminishes as sample sizes increase, which is in line with Roemer et al.'s [77] findings. Hence, HTMT and HTMT+ can be considered slightly more conservative, whereas HTMT2 and HTMT2+ can be considered more liberal metrics for discriminant validity assessment.

Finally, Table A7 shows the results for combinations of negative and positive loadings when the constructs are 0.8 correlated. The findings parallel those obtained in the setting with 0.3 construct correlations. HTMT and HTMT2 produce substantial amounts of inadmissible solutions, while their absolute variants (HTMT+ and HTMT2+) yield almost the same values as in the situation where all loadings are positive.[6]

To summarize, we find that particularly HTMT2 suffers from inadmissible solutions, not only in scenarios where negative loadings occur systematically (whether expected or not), but also in situations where construct correlations are low. While using HTMT2+ resolves this issue, the metric's performance doesn't substantially improve on the HTMT+ metric, even when loadings are heterogeneous (i.e., in this situation HTMT+ has a slight upwards bias while HTMT2+ has a downwards bias for smaller sample sizes). When loadings are homogeneous, HTMT+ actually shows a slightly better performance. In light of the direction of the bias, HTMT2+ can be considered more liberal as the metric shows a slight underestimation tendency. This behaviour can be problematic when comparing the metric against a firm cutoff value. Considering the occurrence of inadmissible solutions and the magnitude and size of the biases produced by the metrics, our results therefore suggest that researchers should primarily rely on the HTMT+ in their discriminant validity assessment tasks. However, they may also consider HTMT2+, especially when the

[6] Compared with Table A.6, the slight deviations in HTMT2+ results at the third decimal place are due to rounding.

Table A7
Simulation results (construct correlation 0.8, negative and positive loadings).

Sample size	HTMT	Number of inadmissible solutions	HTMT+	Number of inadmissible solutions	HTMT2	Number of inadmissible solutions	HTMT2+	Number of inadmissible solutions
Expected HTMT value: 0.8; first loading of the first construct -0.5, all other loadings 0.5								
100	1.860	5054	0.821	0	0.930	9809	0.775	0
200	2.178	5034	0.804	0	0.973	9997	0.779	0
300	2.331	4967	0.804	0	n/a	10,000	0.789	0
10,000	5.869	5067	0.800	0	n/a	10,000	0.800	0
Expected HTMT value: 0.8; first loading of the first construct -0.7, all other loadings 0.7								
Sample size	HTMT	Number of inadmissible solutions	HTMT+	Number of inadmissible solutions	HTMT2	Number of inadmissible solutions	HTMT2+	Number of inadmissible solutions
100	2.969	5139	0.801	0	n/a	10,000	0.793	0
200	3.417	5104	0.799	0	n/a	10,000	0.795	0
300	4.031	5071	0.800	0	n/a	10,000	0.797	0
10,000	9.013	5019	0.800	0	n/a	10,000	0.800	0
Expected HTMT value: 0.8; loadings: -0.9 / 0.8 / 0.7 / 0.6 (first construct), 0.9 / 0.8 / 0.7 / 0.6 (second construct)								
Sample size	HTMT	Number of inadmissible solutions	HTMT+	Number of inadmissible solutions	HTMT2	Number of inadmissible solutions	HTMT2+	Number of inadmissible solutions
100	n/a	10,000	0.806	0	n/a	10,000	0.793	0
200	n/a	10,000	0.805	0	n/a	10,000	0.795	0
300	n/a	10,000	0.806	0	n/a	10,000	0.798	0
10,000	n/a	10,000	0.806	0	n/a	10,000	0.800	0

loadings pattern is quite heterogenous (e.g., 0.95/0.8/0.65), and the sample is not too small (e.g., >300).

Due to the nature of this article, we focused on a limited number of simulation factors, which, however (1) directly relate to the reviewer's concern, and (2) extend the initial design to accommodate more realistic settings. Future research may add to these findings by researching the impact of further design factors on the metrics' behavior, such as the number of indicators, which partly dictates the degree of possible heterogeneity in indicator loadings—assuming a certain quality of measurement models. Finally, future research could also extend on the concept of discriminant validity testing as implied by HTMT and consider measurement congruence, as recently discussed by [115].

CRediT Author Statement

Christian M. Ringle: Conceptualization, Methodology, Resources, Software, Writing – original draft, Writing – review & editing, Visualization; **Marko Sarstedt:** Conceptualization, Methodology, Resources, Writing – review & editing, Visualization; **Noemi Sinkovics:** Conceptualization, Methodology, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration; **Rudolf R. Sinkovics:** Conceptualization, Resources, Writing – original draft, Writing – review & editing.

References

[1] H. Wold, Path models with latent variables: the NIPALS approach, in: H.M. Blalock, A. Aganbegian, F.M. Borodkin, R. Boudon, V. Capecchi (Eds.), Quantitative Sociology: International Perspectives on Mathematical and Statistical Modeling, Academic Press, New York, 1975, pp. 307–357.
[2] J.-B. Lohmöller, Latent Variable Path Modeling with Partial Least Squares, Physica, Heidelberg, 1989, doi:10.1007/978-3-642-52512-4.

[3] M. Tenenhaus, V. Esposito Vinzi, Y.-M. Chatelin, C. Lauro, PLS Path Modeling, Computational Statistics & Data Analysis 48 (1) (2005) 159–205, doi:10.1016/j.csda.2004.03.005.
[4] V. Esposito Vinzi, L. Trinchera, S. Amato, PLS path modeling: from foundations to recent developments and open issues for model assessment and improvement, in: V. Esposito Vinzi, W.W. Chin, J. Henseler, H. Wang (Eds.), Handbook of Partial Least Squares: Concepts, Methods and Applications (Springer Handbooks of Computational Statistics Series, vol. II), Springer, Heidelberg, Dordrecht, London, New York, 2010, pp. 47–82, doi:10.1007/978-3-540-32827-8_3.
[5] W.W. Chin, The partial least squares approach to structural equation modeling, in: G.A. Marcoulides (Ed.), Modern Methods for Business Research, Lawrence Erlbaum, Mahwah, NJ, 1998, pp. 295–358.
[6] J.F. Hair, C.M. Ringle, M. Sarstedt, PLS-SEM: indeed a silver bullet, Journal of Marketing Theory and Practice 19 (2) (2011) 139–151, doi:10.2753/mtp1069-6679190202.
[7] M. Sarstedt, J.F. Hair, C.M. Ringle, “PLS-SEM: indeed a silver bullet” – retrospective observations and recent advances, Journal of Marketing Theory & Practice (2022) Advance online publication, doi:10.1080/10696679.2022.2056488.
[8] T. Köhler, J.M. Cortina, Play it again, Sam! An analysis of constructive replication in the organizational sciences, Journal of Management 47 (2) (2021) 488–518, doi:10.1177/0149206319843985.
[9] R.D. Peng, Reproducible research in computational science, Science 334 (6060) (2011) 1226–1227, doi:10.1126/science.1213847.
[10] LERUOpen science and its role in universities: a roadmap for cultural change, League of European Research Universities, Leuven, 2018.
[11] Fosterscience, What is Open Science? Four pillars of Open Science: putting OS into practice. 2018 https://www.fosteropenscience.eu/learning/what-is-open-science/#/id/5ab8ea32dd1827131b90e3ac Accessed February 25, 2023.
[12] J. Cribb, T.S. Hartomo, Open science sharing knowledge in the global century, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Collingwood, 2010.
[13] C.M. Castille, L.M. Kreamer, B.H. Albritton, G.C. Banks, S.G. Rogelberg, The open science challenge: adopt one practice that enacts widely shared values, Journal of Business and Psychology 37 (3) (2022) 459–467, doi:10.1007/s10869-022-09806-2.
[14] G.S. Christensen, J. Freese, E. Miguel, Transparent and reproducible social science research: how to do open science, University of California Press, Oakland, CA, 2019, doi:10.2307/j.ctvpb3xkg.
[15] P.G. Hensel, Dissecting the tension of open science standards implementation in management and organization journals, Accountability in Research 20 (3) (2021) 150–175, doi:10.1080/08989621.2021.1981870.
[16] E. Ciavolino, M. Aria, J.-H. Cheah, J.L. Roldán, A tale of PLS structural equation modelling: episode I - a bibliometrix citation analysis, Social Indicators Research 164 (3) (2022) 1323–1348, doi:10.1007/s11205-022-02994-7.
[17] R.L. Daft, Why I Recommended That Your Manuscript Be Rejected and What You Can Do About It, in: L.L. Cummings, P.J. Frost (Eds.), Publishing in the organizational sciences, R.D. Irwin, Homewood, IL, 1985, pp. 164–182.
[18] W.O. Bearden, R.G. Netemeyer, K.L. Haws, Handbook of marketing scales: multi-item measures for marketing and consumer behavior research, 3rd ed., Sage, Thousand Oaks, CA, 2011.
[19] G.C. Bruner, Marketing scales handbook: multi-item measures for consumer insight research, GCBII Productions, Fort Worth, TX, 2021.
[20] N. Coviello, H. Yli-Renko, Handbook of Measures for International Entrepreneurship Research: Multi-item Scales Crossing Disciplines and Contexts, Edward Elgar, Cheltenham, 2016, doi:10.4337/9781784711405.
[21] J.L. Price, Handbook of organizational measurement, International Journal of Manpower 18 (4/5/6) (1997) 305–558, doi:10.1108/01437729710182260.
[22] F.J. Smith, Organizational surveys: The diagnosis and betterment of organizations through their members, Psychology Press, London, 2003, doi:10.4324/9781410607270.
[23] M. Appelbaum, H. Cooper, R.B. Kline, E. Mayo-Wilson, A.M. Nezu, S.M. Rao, Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report, American Psychologist 73 (2018) 3–25, doi:10.1037/amp0000191.
[24] CABS, Academic journal guide. 2021 https://charteredabs.org/academic-journal-guide-2021/. Accessed June 24, 2021.
[25] H. Tüselmann, R.R. Sinkovics, G. Pishchulov, Revisiting the standing of International Business journals in the competitive landscape, Journal of World Business 51 (4) (2016) 487–498, doi:10.1016/j.jwb.2016.01.006.
[26] A.-W. Harzing, Journal Quality List. 2022 https://harzing.com/resources/journal-quality-list. Accessed October 21, 2022.
[27] N. Sinkovics, J. Kim, R.R. Sinkovics, Business-civil society collaborations in South Korea: a multi-stage pattern matching study, Management International Review 62 (4) (2022) 471–516, doi:10.1007/s11575-022-00476-z.
[28] E.E. Rigdon, Rethinking partial least squares path modeling: in praise of simple methods, Long Range Planning 45 (5) (2012) 341–358, doi:10.1016/j.lrp.2012.09.010.
[29] M. Sarstedt, J.F. Hair, C.M. Ringle, K.O. Thiele, S.P. Gudergan, Estimation issues with PLS and CBSEM: Where the bias lies!, Journal of Business Research 69 (10) (2016) 3998–4010, doi:10.1016/j.jbusres.2016.06.007.
[30] J.F. Hair, M. Sarstedt, Explanation plus prediction—the logical focus of project management research, Project Management Journal 52 (4) (2021) 319–322, doi:10.1177/8756972821999945.
[31] B.D. Liengaard, P.N. Sharma, G.T.M. Hult, M.B. Jensen, M. Sarstedt, J.F. Hair, C.M. Ringle, Prediction: coveted, yet forsaken? Introducing a cross-validated predictive ability test in partial least squares path modeling, Decision Sciences 52 (2) (2021) 362–392, doi:10.1111/deci.12445.
[32] M. Sarstedt, N.P. Danks, Prediction in HRM research–a gap between rhetoric and reality, Human Resource Management Journal 32 (2) (2022) 485–513, doi:10.1111/1748-8583.12400.
[33] S. Gregor, The nature of theory in information systems, MIS Quarterly 30 (3) (2006) 611–642, doi:10.2307/25148742.

[34] T.K. Dijkstra, PLS' Janus face–response to professor Rigdon's 'rethinking partial least squares modeling: in praise of simple methods, Long Range Planning 47 (3) (2014) 146–153, doi:10.1016/j.lrp.2014.02.004.
[35] T.K. Dijkstra, K. Schermelleh-Engel, Consistent partial least squares for nonlinear structural equation models, Psychometrika 79 (4) (2014) 585–604, doi:10.1007/s11336-013-9370-0.
[36] P.M. Bentler, W. Huang, On components, latent variables, PLS and simple methods: reactions to Rigdon's rethinking of PLS, Long Range Planning 47 (3) (2014) 138–145, doi:10.1016/j.lrp.2014.02.005.
[37] W. Huang, PLSe: Efficient Estimators and Tests for Partial Least Square (Dissertation), University of California Los Angeles, Los Angeles, CA, 2013.
[38] K.-H. Yuan, Y. Wen, J. Tang, Regression analysis with latent variables by partial least squares and four other composite scores: consistency, bias and correction, Structural Equation Modeling: A Multidisciplinary Journal 27 (3) (2020) 333–350, doi:10.1080/10705511.2019.1647107.
[39] E.E. Rigdon, M. Sarstedt, C.M. Ringle, On Comparing Results from CB-SEM and PLS-SEM. Five Perspectives and Five Recommendations, Marketing ZFP 39 (3) (2017) 4–16.
[40] G. Cho, M. Sarstedt, H. Hwang, Comparison of covariance structure analysis, partial least squares path modeling and generalized structured component analysis in factor- and composite models, British Journal of Mathematical and Statistical Psychology 75 (2) (2022) 220–251, doi:10.1111/bmsp.12255.
[41] J.F. Hair, G.T.M. Hult, C.M. Ringle, M. Sarstedt, K.O. Thiele, Mirror, mirror on the wall: a comparative evaluation of composite-based structural equation modeling methods, Journal of the Academy of Marketing Science 45 (5) (2017) 616–632, doi:10.1007/s11747-017-0517-x.
[42] H. Wold, Soft modeling: the basic design and some extensions, in: K.G. Jöreskog, H. Wold (Eds.), Systems Under Indirect Observations: Part II, North-Holland, Amsterdam, 1982, pp. 1–54.
[43] J.F. Hair, G.T.M. Hult, C.M. Ringle, M. Sarstedt, A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM), 3rd ed., Sage, Thousand Oaks, CA, 2022.
[44] G. Russolillo, Non-metric partial least squares, Electronic Journal of Statistics 6 (none) (2012) 1641–1669 29, doi:10.1214/12-EJS724.
[45] G. Cantaluppi, G. Boari, A partial least squares algorithm handling ordinal variables, in: H. Abdi, V.Esposito Vinzi, G. Russolillo, G. Saporta, L. Trinchera (Eds.), The Multiple Facets of Partial Least Squares and Related Methods: PLS, Paris, France, 2014, Springer International Publishing, Cham, 2016, pp. 295–306, doi:10.1007/978-3-319-40643-5_22.
[46] F. Schuberth, J. Henseler, T.K. Dijkstra, Partial least squares path modeling using ordinal categorical indicators, Quality & Quantity 52 (1) (2018) 9–35, doi:10.1007/s11135-016-0401-7.
[47] J.F. Hair, C.M. Ringle, S.P. Gudergan, A. Fischer, C. Nitzl, C. Menictas, Partial least squares structural equation modeling-based discrete choice modeling: an illustration in modeling retailer choice, Business Research 12 (1) (2019) 115–142, doi:10.1007/s40685-018-0072-4.
[48] J.-M. Becker, J.-H. Cheah, R. Gholamzade, C.M. Ringle, M. Sarstedt, PLS-SEM's most wanted guidance, International Journal of Contemporary Hospitality Management 35 (1) (2023) 321–346, doi:10.1108/IJCHM-04-2022-0474.
[49] M. Sarstedt, P. Bengart, A.M. Shaltoni, S. Lehmann, The use of sampling methods in advertising research: a gap between theory and practice, International Journal of Advertising 37 (4) (2018) 650–663, doi:10.1080/02650487.2017.1348329.
[50] U.A. Saari, S. Damberg, L. Frömbling, C.M. Ringle, Sustainable consumption behavior of Europeans: the influence of environmental knowledge and risk perception on environmental concern and behavioral intention, Ecological Economics 189 (2021) 107155, doi:10.1016/j.ecolecon.2021.107155.
[51] J.-M. Becker, I.R. Ismail, Accounting for sampling weights in PLS path modeling: simulations and empirical examples, European Management Journal 34 (6) (2016) 606–617, doi:10.1016/j.emj.2016.06.009.
[52] J.-H. Cheah, J.L. Roldán, E. Ciavolino, H. Ting, T. Ramayah, Sampling weight adjustments in partial least squares structural equation modeling: guidelines and illustrations, Total Quality Management & Business Excellence 32 (13-14) (2021) 1594–1613, doi:10.1080/14783363.2020.1754125.
[53] M.A. Webb, J.P. Tangney, Too good to be true: bots and bad data from Mechanical Turk, Perspectives on Psychological Science (2022), doi:10.1177/17456916221120027.
[54] J. Henseler, C.M. Ringle, R.R. Sinkovics, The use of partial least squares path modeling in international marketing, in: R.R. Sinkovics, P.N. Ghauri (Eds.), Advances in International Marketing, Emerald Bingley, 2009, pp. 277–320, doi:10.1108/S1474-7979(2009)0000020014.
[55] M. Sarstedt, C.M. Ringle, D. Smith, R. Reams, J.F. Hair, Partial least squares structural equation modeling (PLS-SEM): a useful tool for family business researchers, Journal of Family Business Strategy 5 (1) (2014) 105–115, doi:10.1016/j.jfbs.2014.01.002.
[56] N. Kock, P. Hadaya, Minimum sample size estimation in PLS-SEM: the inverse square root and gamma-exponential methods, Information Systems Journal 28 (1) (2018) 227–261, doi:10.1111/isj.12131.
[57] J. Cohen, A power primer, Psychological Bulletin 112 (1) (1992) 155–159, doi:10.1037//0033-2909.112.1.155.
[58] K.J. Rothman, S. Greenland, Planning study size based on precision rather than power, Epidemiology 29 (5) (2018) 599–603, doi:10.1097/ede.0000000000000876.
[59] R.L. Wasserstein, A.L. Schirm, N.A. Lazar, Moving to a world beyond "$p < 0.05$", The American Statistician 73 (sup1) (2019) 1–19, doi:10.1080/00031305.2019.1583913.
[60] J.F. Hair, W.C. Black, B.J. Babin, R.E. Anderson, Multivariate Data Analysis, 8 ed., Cengage Learning, London, 2018.
[61] H. Wang, S. Lu, Y. Liu, Missing data imputation in PLS-SEM, Quality & Quantity 56 (6) (2022) 4777–4795, doi:10.1007/s11135-022-01338-4.
[62] M. Sarstedt, E.A. Mooi, A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistics, 3 ed., Springer, Berlin, 2019, doi:10.1007/978-3-642-12541-6.
[63] M.I. Aguirre-Urreta, M. Rönkkö, Statistical inference with PLSc using bootstrap confidence intervals, MIS Quarterly 42 (3) (2018) 1001–1020.
[64] C.M. Ringle, S. Wende, J.-M. Becker, SmartPLS 4, SmartPLS, Oststeinbek, 2022.

[65] N. Kock, WarpPLS 8.0 User Manual, ScriptWarp Systems, Laredo, TX, 2022.
[66] R Core TeamR: a language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2022 https://www.R-project.org.
[67] M.E. Rademaker, F. Schuberth, Reference manual of the R package cSEM: composite-based structural equation modeling (version 0.5.0), https://cran.r-project.org/web/packages/cSEM/cSEM.pdf, 2022.
[68] S. Ray, N.P. Danks, A. Calero Valdez, R package SEMinR: domain-specific language for building and estimating structural equation models (Version 2.3.2), https://cran.r-project.org/web/packages/seminr/2022.
[69] M. Sarstedt, J.F. Hair, M. Pick, B.D. Liengaard, L. Radomir, C.M. Ringle, Progress in partial least squares structural equation modeling use in marketing research in the last decade, Psychology & Marketing 39 (5) (2022) 1035–1064, doi:10.1002/mar.21640.
[70] A.E. Legate, J.F. Hair, J.L. Chretien, J.J. Risher, PLS-SEM: prediction-oriented solutions for HRD researchers, Human Resource Development Quarterly 34 (1) (2023) 91–109, doi:10.1002/hrdq.21466.
[71] M. Sarstedt, C.M. Ringle, J.F. Hair, Partial least squares structural equation modeling, in: C. Homburg, M. Klarmann, A.E. Vomberg (Eds.), Handbook of Market Research, Springer, Cham, 2022, pp. 587–632, doi:10.1007/978-3-319-05542-8_15-2.
[72] G.T.M. Hult, J.F. Hair, D. Proksch, M. Sarstedt, A. Pinkwart, C.M. Ringle, Addressing endogeneity in international marketing applications of partial least squares structural equation modeling, Journal of International Marketing 26 (3) (2018) 1–21, doi:10.1509/jim.17.0151.
[73] C.M. Ringle, M. Sarstedt, Gain more insight from your PLS-SEM results, Industrial Management & Data Systems 116 (9) (2016) 1865–1886, doi:10.1108/IMDS-10-2015-0449.
[74] G. Franke, M. Sarstedt, Heuristics versus statistics in discriminant validity testing: a comparison of four procedures, Internet Research 29 (3) (2019) 430–447, doi:10.1108/IntR-12-2017-0515.
[75] A.M. Farrell, Insufficient discriminant validity: a comment on Bove, Pervan, Beatty, and Shiu (2009), Journal of Business Research 63 (3) (2010) 324–327, doi:10.1016/j.jbusres.2009.05.003.
[76] J. Henseler, C.M. Ringle, M. Sarstedt, A new criterion for assessing discriminant validity in variance-based structural equation modeling, Journal of the Academy of Marketing Science 43 (1) (2015) 115–135, doi:10.1007/s11747-014-0403-8.
[77] E. Roemer, F. Schuberth, J. Henseler, HTMT2–an improved criterion for assessing discriminant validity in structural equation modeling, Industrial Management & Data Systems 121 (12) (2021) 2637–2650, doi:10.1108/IMDS-02-2021-0082.
[78] G. Shmueli, S. Ray, J.M. Velasquez Estrada, S.B. Chatla, The elephant in the room: predictive performance of PLS models, Journal of Business Research 69 (10) (2016) 4552–4564, doi:10.1016/j.jbusres.2016.03.049.
[79] G. Shmueli, M. Sarstedt, J.F. Hair, J.-H. Cheah, H. Ting, S. Vaithilingam, C.M. Ringle, Predictive model assessment in PLS-SEM: guidelines for using PLSpredict, European Journal of Marketing 53 (11) (2019) 2322–2347, doi:10.1108/EJM-02-2019-0189.
[80] P.N. Sharma, B.D. Liengaard, J.F. Hair, M. Sarstedt, C.M. Ringle, Predictive model assessment and selection in composite-based modeling using PLS-SEM: extensions and guidelines for using CVPAT, European Journal of Marketing (2022) ahead-of-print, doi:10.1108/EJM-08-2020-0636.
[81] F. Schuberth, M.E. Rademaker, J. Henseler, Assessing the overall fit of composite models estimated by partial least squares path modeling, European Journal of Marketing (2022) ahead-of-print, doi:10.1108/EJM-08-2020-0586.
[82] W. Chin, J.-H. Cheah, Y. Liu, H. Ting, X.-J. Lim, T.H. Cham, Demystifying the role of causal-predictive modeling using partial least squares structural equation modeling in information systems research, Industrial Management & Data Systems 120 (12) (2020) 2161–2209, doi:10.1108/IMDS-10-2019-0529.
[83] M. de Rooij, J.D. Karch, M. Fokkema, Z. Bakk, B.C. Pratiwi, H. Kelderman, SEM-based out-of-sample pvredictions, Structural Equation Modeling: A Multidisciplinary Journal 30 (1) (2023) 132–148, doi:10.1080/10705511.2022.2061494.
[84] G. Shmueli, To explain or to predict? Statistical Science 25 (3) (2010) 289–310 22, doi:10.1214/10-STS330.
[85] J.F. Hair, M. Sarstedt, C.M. Ringle, S.P. Gudergan, Advanced Issues in Partial Least Squares Structural Equation Modeling (PLS-SEM), 2nd ed., Sage, Thousand Oaks, CA, 2024.
[86] M. Sarstedt, C.M. Ringle, J.-H. Cheah, H. Ting, O.I. Moisescu, L. Radomir, Structural model robustness checks in PLS-SEM, Tourism Economics 26 (4) (2020) 531–554, doi:10.1177/1354816618823921.
[87] B. Szaszi, A. Higney, A. Charlton, A. Gelman, I. Ziano, B. Aczel, D.G. Goldstein, D.S. Yeager, E. Tipton, No reason to expect large and consistent effects of nudge interventions, Proceedings of the National Academy of Sciences 119 (31) (2022) e2200732119, doi:10.1073/pnas.2200732119.
[88] T. Salzberger, R.R. Sinkovics, B.B. Schlegelmilch, Data equivalence in cross-cultural research: a comparison of classical test theory and latent trait theory based approaches, Australasian Marketing Journal 7 (2) (1999) 23–38, doi:10.1016/S1441-3582(99)70213-2.
[89] J. Henseler, C.M. Ringle, M. Sarstedt, Testing measurement invariance of composites using partial least squares, International Marketing Review 33 (3) (2016) 405–431, doi:10.1108/IMR-09-2014-0304.
[90] L.M. Federer, C.W. Belter, D.J. Joubert, A. Livinski, Y.L. Lu, L.N. Snyders, H. Thompson, Data sharing in PLOS ONE: An analysis of Data Availability Statements, PLoS One 13 (5) (2018) e0194768, doi:10.1371/journal.pone.0194768.
[91] C.F. Camerer, A. Dreber, F. Holzmeister, T.-H. Ho, J. Huber, M. Johannesson, M. Kirchler, G. Nave, B.A. Nosek, T. Pfeiffer, A. Altmejd, N. Buttrick, T. Chan, Y. Chen, E. Forsell, A. Gampa, E. Heikensten, L. Hummer, T. Imai, S. Isaksson, D. Manfredi, J. Rose, E.-J. Wagenmakers, H. Wu, Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015, Nature Human Behaviour 2 (9) (2018) 637–644, doi:10.1038/s41562-018-0399-z.
[92] E.E. Rigdon, M. Sarstedt, J.-M. Becker, Quantify uncertainty in behavioral research, Nature Human Behaviour 4 (4) (2020) 329–331, doi:10.1038/s41562-019-0806-0.

[93] E.E. Rigdon, M. Sarstedt, Accounting for uncertainty in the measurement of unobservable marketing phenomena, in: H. Baumgartner, B. Weijters (Eds.), Measurement in Marketing, Emerald, Bingley, 2022, pp. 53–73, doi:10.1108/S1548-643520220000019003.
[94] APAPublication manual of the American Psychological Association, 7th ed., American Psychological Association, Washington, 2020, doi:10.1037/0000165-000.
[95] R. Pears, G.J. Shields, Cite them right: the essential referencing guide, 12th ed., Bloomsbury Academic, London, 2022.
[96] M. Rönkkö, E. Cho, An updated guideline for assessing discriminant validity, Organizational Research Methods 25 (1) (2022) 6–14, doi:10.1177/1094428120968614.
[97] D. McNeish, Thanks coefficient alpha, we'll take it from here, Psychological Methods 23 (3) (2018) 412–433, doi:10.1037/met0000144.
[98] G. Walsh, V.-W. Mitchell, P.R. Jackson, S.E. Beatty, Examining the antecedents and consequences of corporate reputation: a customer perspective, British Journal of Management 20 (2) (2009) 187–203, doi:10.1111/j.1467-8551.2007.00557.x.
[99] M.A. Jones, D.L. Mothersbaugh, S.E. Beatty, Switching barriers and repurchase intentions in services, Journal of Retailing 76 (2) (2000) 259–274, doi:10.1016/S0022-4359(00)00024-5.
[100] A. Usakli, K.G. Kucukergin, Using partial least squares structural equation modeling in hospitality and tourism, International Journal of Contemporary Hospitality Management 30 (11) (2018) 3462–3512, doi:10.1108/IJCHM-11-2017-0753.
[101] E. Bayonne, J.A. Marin-Garcia, R. Alfalla-Luque, Partial least squares (PLS) in operations management research: insights from a systematic literature review, Journal of Industrial Engineering and Management 13 (3) (2020), doi:10.3926/jiem.3416.
[102] N. Zeng, Y. Liu, P. Gong, M. Hertogh, M. König, Do right PLS and do PLS right: a critical review of the application of PLS-SEM in construction management research, Frontiers of Engineering Management 8 (3) (2021) 356–369, doi:10.1007/s42524-021-0153-5.
[103] H.-M. Lin, M.-H. Lee, J.-C. Liang, H.-Y. Chang, P. Huang, C.-C. Tsai, A review of using partial least square structural equation modeling in e-learning research, British Journal of Educational Technology 51 (4) (2020) 1354–1372, doi:10.1111/bjet.12890.
[104] C.M. Ringle, S. Wende, J.-M. Becker, SmartPLS 3, SmartPLS, Bönningstedt, 2015.
[105] E.R. Thompson, Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS), Journal of Cross-Cultural Psychology 38 (2) (2007) 227–242, doi:10.1177/0022022106297301.
[106] M. van der Gaag, T. Hoffman, M. Remijsen, R. Hijman, L. de Haan, B. van Meijel, P.N. van Harten, L. Valmaggia, M. de Hert, A. Cuijpers, D. Wiersma, The five-factor model of the positive and negative syndrome scale II: a ten-fold cross-validation of a revised model, Schizophrenia Research 85 (1) (2006) 280–287, doi:10.1016/j.schres.2006.03.021.
[107] P. Bech, M. Fava, M.H. Trivedi, S.R. Wisniewski, A.J. Rush, Factor structure and dimensionality of the two depression scales in STAR*D using level 1 datasets, Journal of Affective Disorders 132 (3) (2011) 396–400, doi:10.1016/j.jad.2011.03.011.
[108] J. He, F.J.R. van de Vijver, A general response style factor: Evidence from a multi-ethnic study in the Netherlands, Personality and Individual Differences 55 (7) (2013) 794–800, doi:10.1016/j.paid.2013.06.017.
[109] J. He, D. Bartram, I. Inceoglu, F.J.R. van de Vijver, Response styles and personality traits: a multilevel analysis, Journal of Cross-Cultural Psychology 45 (7) (2014) 1028–1045, doi:10.1177/0022022114534773.
[110] P. De Boeck, F. Rijmen, Response times in cognitive tests: interpretation and importance, in: M.J. Margolis, R.A. Feinberg (Eds.), Integrating Timing Considerations to Improve Testing Practices, Routledge, New York, NY, 2020, pp. 142–149.
[111] J.P. Peter, Reliability: a review of psychometric basics and recent marketing practices, Journal of Marketing Research 16 (1) (1979) 6–17, doi:10.2307/3150868.
[112] D. Folashade, R Package foreach: provides foreach looping construct (version 1.5.2), https://cran.r-project.org/web/packages/foreach/2022.
[113] J.W. Osborne, Best Practices in Data Cleaning: A Complete Guide to Everything You Need to Do Before and After Collecting Your Data, Sage, Thousand Oaks, CA, 2013.
[114] T.K. Dijkstra, J. Henseler. Consistent partial least squares path modeling, MIS Quarterly 39 (2) (2015) 297-316. doi:10.25300/MISQ/2015/39.2.02.
[115] G.R. Franke, M. Sarstedt, N.P. Danks, Assessing measurement congruence in nomological networks, Journal of Business Research 130 (2021) 318–334, doi:10.1016/j.jbusres.2021.03.003.