Received: 21 March 2024 | Accepted: 21 March 2024

DOI: 10.1111/jpim.12738

CATALYST

JOURNAL OF PRODUCT INNOVATION MANAGEMENT

Same model, same data, but different outcomes: Evaluating the impact of method choices in structural equation modeling

Marko Sarstedt[1,2] | Susanne J. Adler[1] | Christian M. Ringle[3,4] | Gyeongcheol Cho[5] | Adamantios Diamantopoulos[6] | Heungsun Hwang[7] | Benjamin D. Liengaard[8]

[1]Institute for Marketing, Ludwig-Maximilians-University Munich, Munich, Germany

[2]Faculty of Economics and Business Administration, Babeş-Bolyai University, Cluj-Napoca, Romania

[3]Institute of Management and Decision Sciences, Hamburg University of Technology, Hamburg, Germany

[4]College of Business, Law and Governance, James Cook University, Townsville, Queensland, Australia

[5]Department of Psychology, The Ohio State University, Columbus, Ohio, USA

[6]Department of Marketing and International Business, University of Vienna, Vienna, Austria

[7]Department of Psychology, McGill University, Montreal, Quebec, Canada

[8]Department of Economics and Business Economics, Aarhus University, Aarhus, Denmark

Correspondence
Marko Sarstedt, Institute for Marketing, Ludwig-Maximilians-University Munich, Munich, Germany; Faculty of Economics and Business Administration, Babeş-Bolyai University, Cluj-Napoca, Romania.
Email: sarstedt@lmu.de

Abstract

Scientific research demands robust findings, yet variability in results persists due to researchers' decisions in data analysis. Despite strict adherence to state-of-the-art methodological norms, research results can vary when analyzing the same data. This article aims to explore this variability by examining the impact of researchers' analytical decisions when using different approaches to structural equation modeling (SEM), a widely used method in innovation management to estimate cause–effect relationships between constructs and their indicator variables. For this purpose, we invited SEM experts to estimate a model on absorptive capacity's impact on organizational innovation and performance using different SEM estimators. The results show considerable variability in effect sizes and significance levels, depending on the researchers' analytical choices. Our research underscores the necessity of transparent analytical decisions, urging researchers to acknowledge their results' uncertainty, to implement robustness checks, and to document the results from different analytical workflows. Based on our findings, we provide recommendations and guidelines on how to address results variability. Our findings, conclusions, and recommendations aim to enhance research validity and reproducibility in innovation management, providing actionable and valuable insights for improved future research practices that lead to solid practical recommendations.

KEYWORDS

metascience, scientific transparency, structural equation modeling, uncertainty

1 | INTRODUCTION

To address contemporary grand challenges and to identify meaningful managerial and policy implications, research in any field—including innovation management—has to produce objective, reliable, and valid results. Subjectivity, however, plays a crucial role in scientific studies as each stage of the scientific process comes with numerous decisions that may translate into different analytical outcomes (e.g., Buchanan et al., 1998; Gelman & Loken, 2014; Schweinsberg et al., 2021; Wagenmakers et al., 2022). Even under rigid adherence to the scientific method, high ethical standards, and state-of-the-art approaches for maximizing reproducibility, research results are variable (Breznau et al., 2022). A potential reason for the observed variability in results lies in the complexity and ambiguity inherent in the process of data analysis, due to, for example, researchers' choice of measures and data treatment options, model building activities, and selection and use of statistical estimators (Simmons et al., 2011).

Researchers in neuroscience, psychology, and sociology have started examining how researchers' degrees of freedom in statistical designs translate into different results and study implications. For example, Silberzahn et al. (2018) recruited 29 teams of researchers with strong statistical backgrounds and asked them to answer the same research question ("Are football referees more likely to give red cards to players with dark skin than to players with light skin?") with the same dataset. Similarly, Botvinik-Nezer et al. (2020) invited 70 independent teams to test nine hypotheses on a single neuroimaging dataset. Finally, in Breznau et al. (2022) 73 research teams used the same cross-country dataset to test the hypothesis that more immigration will reduce public support for government provision of social policies. Across all three studies, the researchers obtained highly variable, sometimes divergent results, resulting from the application of different data treatment options, the specification of different models, and the use of a wide array of analytical techniques. While these studies highlight the impact of researchers' manifold analytical choices on the results and implications, scant attention has been paid to evaluating whether different researchers arrive at the same or at least similar findings when analyzing the *same* theoretical model using *identical* data but *alternative* methods.

We seek to explore this issue in the methodological field of structural equation modeling (SEM), which—as our review of major innovation management journals will show—features prominently in innovation management for estimating cause-effect relationships between constructs and their indicator variables. Its ability to model complex interrelationships between multiple layers of constructs, while simultaneously accounting for measurement error inherent in the indicators (Hair et al., 2022, chap. 1) makes SEM useful for a plethora of research questions in the innovation management field, which routinely considers relationships between unobservable phenomena such as innovation orientation (Siguaw et al., 2006; Stock & Zacharias, 2011), trust in business partners (Jean et al., 2014; Kemper et al., 2013), and user-centric design capabilities (Cautela et al., 2022).

Practitioner points

- Managers rely on academic research to make informed decisions in their everyday operations. Media outlets often portray research results as unambiguous without acknowledging their uncertainty, which may trigger erroneous conclusions.
- By showcasing the results' variability when analyzing the same model with the same dataset, our study underlines the ambiguity that comes with any statistical analysis—even under relatively controlled conditions.
- Companies should assume different analytical perspectives when working with data in order to safeguard the robustness of the implications drawn from any analysis.

To estimate structural equation models, researchers can draw on a variety of methods, which differ in terms of how they statistically approximate constructs and in their optimization routines (Cho et al., 2022). Because of these differences, the choice of an SEM method inevitably comes with explicit or implicit assumptions regarding the phenomena under study, our ability to measure them comprehensively, and the best means to estimate relations between them (e.g., Rigdon et al., 2017). Numerous simulation studies have evaluated the efficacy of SEM methods from different perspectives and across various data and model constellations in an effort to pinpoint situations in which each method stands out—or falls short of expectations (e.g., Cho et al., 2023; Hair et al., 2017; Reinartz et al., 2009; Sarstedt et al., 2016). While these simulation studies show that neither of the methods is universally superior across all constellations, scant attention has been devoted to the question as to whether applications of these methods by different researchers converge on the same results (e.g., in terms of significance) and findings (e.g., in terms of managerial implications).

The application of any SEM method comes with various degrees of freedom, even when applied to a single model and dataset. For example, estimating a model using one method may identify certain indicators as unreliable, prompting the researcher to modify the model set-up, while another method may not produce such issues. Similarly, whereas one researcher may put strong emphasis on achieving model fit and therefore adjusts the model, for example, on the grounds of model modification indices (Diamantopoulos & Siguaw, 2000; Rigdon, 1998), another researcher may focus more on the model's predictive power whose maximization may require no such modifications, despite low overall model fit. Such differences in analytical workflows may yield different estimates that in turn may entail divergent interpretations and implications for follow-up research, managerial practice, and policy advice. Simulation studies are uninformative in this regard as they do not account for the human element in the data analysis workflow. Specifically, such studies are designed to produce strong discrepancies between method performances to highlight their advantages and disadvantages in selected data and model constellations (e.g., Paxton et al., 2001).

Addressing these concerns, we invited leading experts in the SEM field to estimate a prespecified model on absorptive capacity's impact on organizational innovation and performance (Ali et al., 2016) using the same data, but drawing on different SEM estimators that they developed or mastered in their careers. These experts were asked to analyze the model and disclose their analytical workflows, which may entail adjusting the model to the algorithmic requirements or model estimates. By focusing on the researchers' model estimation choices, we highlight one major source of results' variability (Rigdon et al., 2020), thereby extending prior research which allowed researchers with maximum degrees of freedom in terms of initial model choice (Botvinik-Nezer et al., 2020; Breznau et al., 2022; Silberzahn et al., 2018). If such analyses yield similar results, researchers can speak with one voice on an issue. Alternatively, the estimated effects may be highly contingent on analysis strategies. If so, then subjectivity in data analysis workflows and the ensuing ambiguity in scientific results can be made transparent.

Our results show that while the SEM methods produce estimates that are well within each other's confidence regions, their sizes and sometimes also their significance levels vary considerably, depending on the experts' analytical paths taken. We also find that the experts' workflows differ considerably, for example regarding their decisions on whether or not to modify the model, implying that the research articles that would emerge from each workflow may differ substantially.

By showcasing the impact of analytical decisions on research results, our study makes a case for making every step of the analysis transparent—even in settings where researchers have comparably few degrees of freedom. Our results also suggest that researchers should acknowledge the uncertainty that comes with any statistical analysis and account for alternative analytical workflows (e.g., via robustness checks). Based on our findings, we provide recommendations and guidelines on how to address uncertainty and results variability.

2 | ANALYTICAL APPROACH

2.1 | Structural equation modeling in innovation research and related fields

There are broadly two approaches for estimating structural equation models: Factor-based and component-based SEM (e.g., Rigdon et al., 2017).[1] In factor-based SEM, as carried out by software programs such as AMOS, LISREL, or Mplus, the constructs are represented as common factors, which implies that a construct is an external reality independent of observed variables, "causing" them to covary (Jöreskog, 1978). In estimating the model parameters, the method therefore draws on the indicators' common variance, assuming that it can be fully explained as a function of the construct (the common factor) plus its (unique) error variance (Hair et al., 2017). In contrast, component-based SEM relies on weighted sums of observed variables (i.e., components) to approximate constructs (Tenenhaus, 2008). Using a component implies that a construct is an aggregation of observed variables that acts as if it were a unidimensional entity. Model estimation in component-based SEM, therefore, does not focus on the common variance but considers the indicators' total variance in the model estimation (Sarstedt et al., 2016). Partial least squares (PLS; Lohmöller, 1989; Wold, 1982) and generalized structured component analysis (GSCA; Hwang & Takane, 2004, 2014) are well-developed and widespread methods for component-based SEM (Hwang et al., 2020).[2]

Research has also brought forward various methods that seek to bridge these two SEM domains by adjusting the component-based estimates to conform with a common factor model, including GSCA_M (Hwang et al., 2017), integrated GSCA (Hwang et al., 2021), and consistent PLS (PLSc; Dijkstra, 2010; Dijkstra & Schermelleh-Engel, 2014). Among these methods, particularly PLSc has gained popularity among researchers with applications spread across various business research domains, including innovation management (e.g., Berndt et al., 2023; Suder et al., 2022; Wiesböck et al., 2020).

To analyze the prevalence of the various SEM estimators in innovation research, we conducted a systematic literature review of major journals in the field. Our literature review covers the following nine innovation management and general management journals that frequently publish research in this field: *Academy of Management Journal*, *Creativity and Innovation*

[1]In line with, for instance, with Hwang et al. (2020), we use the terms composites and components interchangeably in this research.

[2]We do not discuss the methods in detail here, but refer to the extensive body of literature that provides their technical underpinnings.

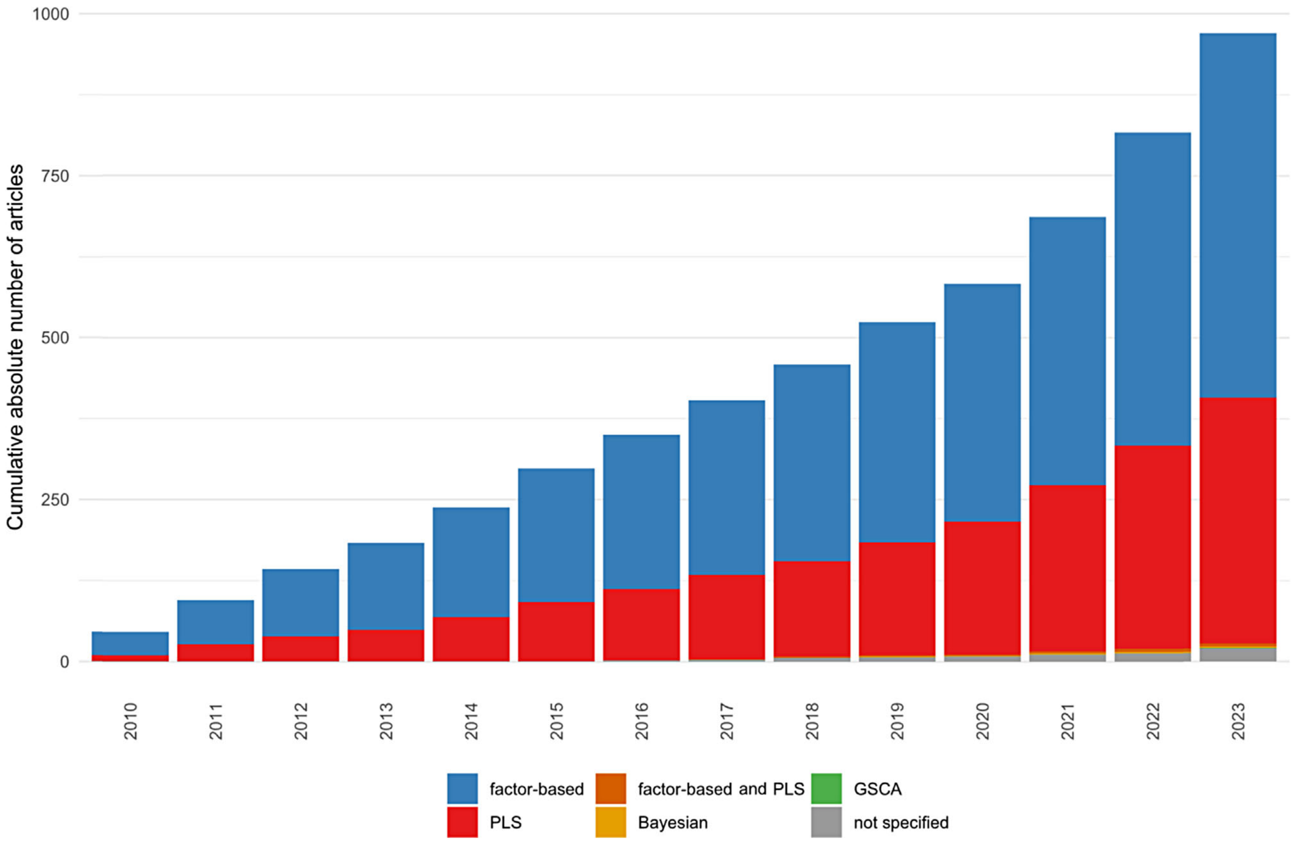

FIGURE 1 Cumulative number of articles using SEM per year and SEM type.

Management, *Journal of Marketing*, *Journal of Product Innovation Management*, *R & D Management*, *Research Policy*, *Strategic Management Journal*, *Technological Forecasting and Social Change*, and *Technovation*. Our analysis of all articles published between 2010 and 2023 demonstrates the relevance of SEM for innovation management research. We find that of the 970 studies that applied SEM, 562 relied on factor-based SEM, while 380 articles used component-based PLS. Five articles applied both methods (e.g., Pemartín et al., 2018), two articles applied a Bayesian SEM approach, one article used GSCA, and 20 articles did not specify the SEM type (Figure 1).[3] Our review also shows that component-based SEM use has grown in popularity over the last decade. Specifically, we find that the number of studies using PLS ($b_{\text{quadratic term}} = 0.50$, $p < 0.001$) and factor-based SEM ($b_{\text{quadratic term}} = 0.61$, $p = 0.002$) both exhibit a positive quadratic trend of time, indicating that the methods' use has accelerated over time.

Considering the results of our literature review, the analysis covers (1) maximum likelihood-based covariance-based SEM (CB-SEM) as the default factor-based method as well as (2) PLS and (3) GSCA as different component-based methods. In addition, we consider (4) GSCA_M and (5) PLSc as two novel approaches whose use has started gaining momentum in business research, even though our analysis disclosed only a few applications in innovation management (e.g., Alesanco-Llorente et al., 2023).

The principal investigators (PIs) invited the following leading experts to provide their analyses using methods that they developed or mastered in their careers:[4] Adamantios Diamantopoulos (CB-SEM), Gyeongcheol Cho (GSCA), Heungsun Hwang (GSCA_M), and Benjamin D. Liengaard (PLS and PLSc).[5]

[3]In most cases where researchers did not specify the type, SEM was used as a complementary method to test a subset of hypotheses or for robustness checks.

[4]The PIs are Marko Sarstedt, Susanne J. Adler, and Christian M. Ringle. Adamantios Diamantopoulos has published a textbook (Diamantopoulos & Siguaw, 2000) on and numerous Scopus Q1 articles using factor-based SEM. Gyeongcheol Cho and Heungsun Hwang have developed the GSCA method and its extensions; their research has been published in Hwang and Takane (2004) and in various Scopus Q1 articles. Benjamin D. Liengaard has developed methodological extensions of PLS, which have been published in Scopus Q1 articles.

[5]Benjamin D. Liengaard was invited to provide two reports as the PLS analysis merely replicated and extended Ali et al. (2016).

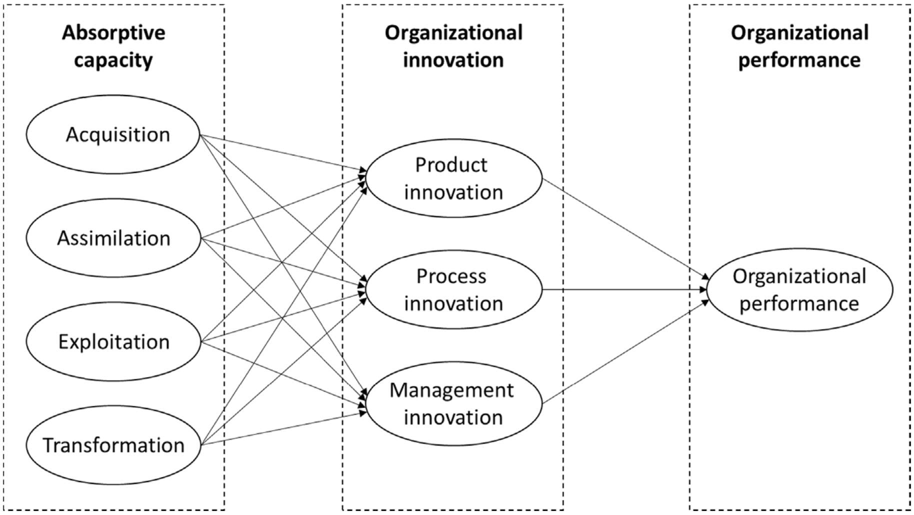

FIGURE 2 Research model (adapted from Ali et al., 2016).

2.2 | Research model and expert briefing

The experts were asked to estimate Ali et al.'s (2016) model on the impact of absorptive capacity (i.e., a firm's "ability to recognize the value of new information, assimilate it, and apply it to commercial ends;" Cohen & Levinthal, 1990, p. 128) on organizational innovation and performance. Specifically, the model considers four dimensions of absorptive capacity, which are hypothesized to impact organizational performance through product, process, and management innovation (Figure 2). Ali et al. (2016) associate higher absorptive capacity with a higher degree of organizational innovation concerning a firm's products, processes, and management operations that in turn increase organizational performance in terms of, for example, market share or profitability. Ali et al. (2016, p. 5318) propose the following hypotheses, which we also examine in this article[6]:

> **Hypothesis 1.** Acquisition relates positively to product innovation, process innovation, and management innovation.

> **Hypothesis 2.** Assimilation relates positively to product innovation, process innovation, and management innovation.

> **Hypothesis 3.** Transformation relates positively to product innovation, process innovation, and management innovation.

> **Hypothesis 4.** Exploitation relates positively to product innovation, process innovation, and management innovation.

> **Hypothesis 5.** Product innovation, process innovation, and management innovation relate positively to organizational performance.

Each construct is operationalized with three to seven items, drawing on a reflective measurement specification (Diamantopoulos & Siguaw, 2006). Ali et al. (2016) used a dataset of $n = 195$ industrial firms in South Korea, employed PLS to test the hypotheses, and found support that acquisition (Hypothesis 1), assimilation (Hypothesis 2), and exploitation (Hypothesis 4) affect organizational performance. On the contrary, the authors do not find that transformation impacts organizational innovation (Hypothesis 3). The results partially support Hypothesis 5 since process and management innovation significantly impact organizational performance, while product innovation does not. The PIs provided the experts with the hypothesized model and the construct operationalizations—as reported by Ali et al. (2016)—and the study's original dataset. This setup closely mirrors a common starting point for a research project's data analysis stage. In the following, each expert estimated the model and provided a complete workflow, documenting the decisions they made during the process, such as model modifications to ensure the goodness-of-fit and increase the model's predictive power. The PIs did however not provide the experts with a structured outline for the data analysis to prevent interfering with their workflow. After collecting the results from each analysis,

[6]Ali et al. (2016) also proposed a sixth hypothesis that assumes a configural perspective of absorptive capacity's impact on organizational performance. As its analysis requires running a fuzzy-set qualitative comparative analysis (Fiss, 2011), we disregard this hypothesis.

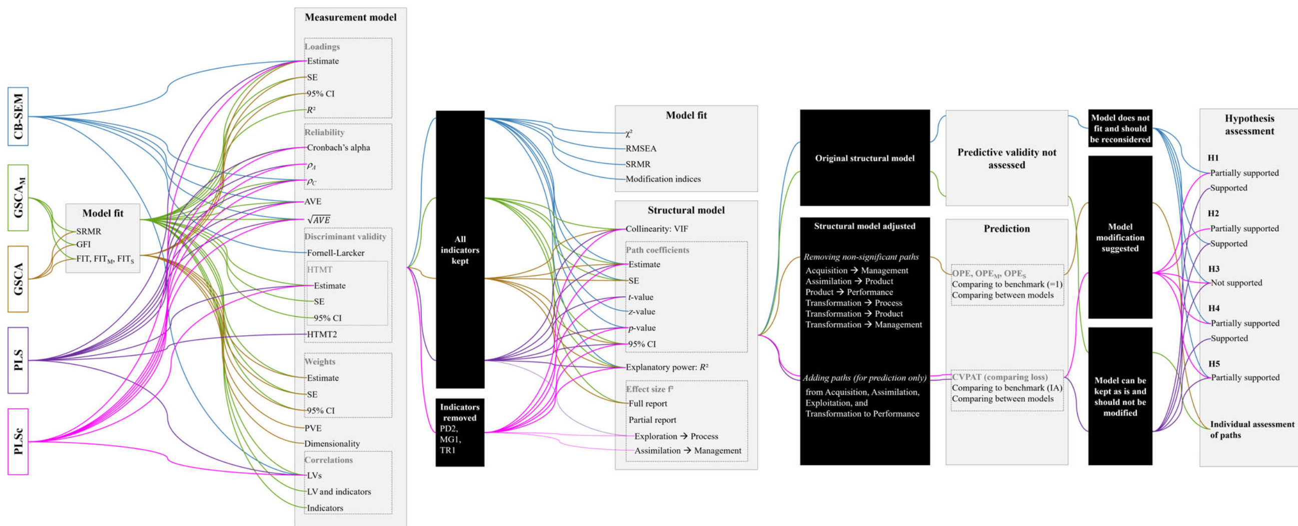

FIGURE 3 Workflow overview.

the PIs examined the experts' workflows to identify which methods and metrics they used and compare the results concerning the hypotheses.

We provide further material on the Open Science Framework (OSF), including details on the literature review's methodology and data, the expert reports, and the original data set from Ali et al. (2016): https://osf.io/29spz/.

3 | RESULTS

We first address the question to what degree the workflows varied among the experts. We then turn our attention to whether the observed workflow differences produced divergent results and findings in terms of the relationships between absorptive capacity, organizational innovation, and organizational performance in the originally hypothesized model.

3.1 | How much did the workflows vary between the experts?

The experts employed a versatility of data analytical workflows, which we visualize in Figure 3. All experts evaluated the quality of the constructs' measurement models. Comparing their approaches to measurement model assessment, we observe some variation in the criteria used, depending on the SEM method. For example, in terms of internal consistency reliability assessment, the CB-SEM report only documents composite reliability, while other experts (additionally) consider Cronbach's alpha (Cronbach, 1951) or ρ_A (Dijkstra & Henseler, 2015). Nevertheless, all metrics generally converge upon supporting the measures' internal consistency reliability, convergent validity, and discriminant validity—with two exceptions. The expert using PLSc noted that the analysis produced low loadings in several indicators that had adverse consequences for the measures' convergent validity and therefore deleted one item in each of the product innovation, management innovation, and transformation constructs. The model estimation using the other methods did not produce such issues. Hence, the PLSc analysis relies on a different measurement model set-up compared with that of the other methods. Another expert applied a confirmatory factor analysis prior to analyzing the structural equation model (Anderson & Gerbing, 1988), bemoaning the measures' poor fit. While the introduction of error covariances and cross-loadings would improve fit, the expert noted that such a step would "violate the principles of unidimensional measurement and the interpretability of the measurement model." All other experts did not consider the measures' fit on the grounds of a confirmatory factor or confirmatory composite analysis (Hair et al., 2020; Jöreskog et al., 2016; Schuberth et al., 2018).

Our analysis of the experts' structural model assessment workflows reveals further differences that are relevant to the results' implications. For example, the analyses using GSCA, $GSCA_M$, and CB-SEM explicitly consider the model's overall fit (Bagozzi & Yi, 1988; Cho et al., 2020), albeit with different results.[7] While the GSCA and $GSCA_M$ analyses indicate an acceptable model

[7]We also note that the experts situated the fit assessment in different elements stages of the process. While the GSCA and $GSCA_M$ analyses initially focused on model fit assessment, the CB-SEM analysis considered overall model fit after the measurement validation.

fit, this is not the case with CB-SEM where the expert explicitly noted that "the empirical results question the specification of product innovation, process innovation, and management innovation as *parallel* mediators, unrelated to each other" and suggested specifying a serial mediation model (product → process → management innovation). In a follow-up exchange with one of the PIs, the expert noted that he did not implement any model modifications due to his lack of substantive knowledge on the topic. On the contrary, the expert applying GSCA altered the model structure, not in response to a potential misfit, but as supplementary evidence for the model's robustness. Specifically, the expert removed 6 of the 15 paths (40%) that the analysis of the original model rendered as nonsignificant, finding that the reduced model exhibits a higher predictive power than the original model. Due to this change in the model specification, the impact of assimilation on management innovation increased from 0.34 to 0.41, thereby inducing a 54% rise in this construct's effect size from $f^2 = 0.13$ to 0.20.

Different from the others, the expert applying PLS and PLSc did not comment on model fit, but solely focused on predictive power analyses using a test based on *k*-fold cross-validation. The predictive power analyses involved comparing the original model's predictive power with that of a saturated model with path relationships between acquisition, assimilation, exploitation, transformation, and organizational performance, which the expert justified on the grounds of prior research.

Finally, all experts document the path coefficient estimates and their significances, but differences emerge in how they evaluated the hypotheses. Specifically, the CB-SEM, PLS, and PLSc reports provide aggregated results on Hypothesis 1 to Hypothesis 5 (i.e., they consider the antecedent constructs jointly), while the GSCA and GSCA$_M$ reports focus on individual paths. This has implications for readers who may perceive the support for a hypothesis differently if it is labeled as "partial support" versus if each path is considered separately.

Consolidating the workflows suggests that individual articles emerging from the various analyses would be very different. Specifically, one expert would not report the model as is (CB-SEM), two experts would report a modified model (GSCA and PLSc), while in the remaining two cases (GSCA$_M$ and PLS), no model modification took place. Other differences include which metrics to report, whether to content with point estimates or to document the estimates' variability—for instance, when reporting loadings (Chin, 1998; Rigdon, 1998) and HTMT values (Henseler et al., 2015)—and whether to conduct a predictive validity assessment (Cho et al., 2023; Sharma et al., 2023).

3.2 | How much did results vary between the experts?

Comparing the structural model results documented in Figure 4, we see a clear pattern. The different SEM methods largely align with the results obtained by Ali et al. (2016). Specifically, all methods find at least partial support for the effects of acquisition (Hypothesis 1), assimilation (Hypothesis 2), and exploitation (Hypothesis 4) on organizational innovation. Furthermore, they find partial support for organizational innovation's impact on organizational performance (Hypothesis 5). None of the analyses supports a significant relation between transformation absorptive capacity and organization innovation (Hypothesis 3).

However, the devil lies in the details. Experts who used SEM methods that assume a common factor model (CB-SEM, GSCA$_M$, and PLSc) documented estimates that were much more variable than those produced by component-based methods, thereby producing divergent implications. For example, according to the PLSc and particularly CB-SEM analyses, management innovation is the strongest driver of organizational performance, followed by process and product innovation. Under GSCA$_M$, however, process innovation has a stronger impact on organizational performance than management innovation, while their impact is practically identical under GSCA and PLS. Considering the methods altogether, we find that the average range across all path relations is 0.097 with a minimum of 0.038 for the relationship between acquisition and product innovation and a maximum of 0.171 for the relationship between management innovation and organizational performance.

In addition, while experts who applied component-based SEM methods reported similar path coefficient estimates, GSCA produced wider confidence intervals, which yielded different conclusions regarding the significance in two of the 15 relations (13.33%) compared with PLS. Specifically, the effects of assimilation on product innovation and acquisition on management innovation are not significant under GSCA, but significant under PLS. This result could lead researchers to consider the significant effect as "present" and the non-significant effect as "absent." Since the methods' path coefficients do not differ much, assuming such a dichotomy would misinform researchers (Rigdon, 2023).

Further divergences emerge regarding the effect sizes of the antecedents of organizational performance. For example, the GSCA$_M$ analysis reports an effect size of exploitation on process innovation ($f^2 = 0.31$), which is about 50% higher than that of GSCA ($f^2 = 0.21$). We find similar divergences for the R^2 values and their interpretation. For example, CB-SEM and PLSc produce R^2 values

FIGURE 4 Structural model results across different methods. The bars represent 95% confidence intervals as reported in the articles. GSCA (original model) refers to the results from the original model specification, while GSCA (revised model) refers to the results after eliminating previously nonsignificant paths.

for process innovation of close to 0.7, while the other methods' values are slightly larger than 0.5. The average R^2 range of all constructs' R^2 values is 0.132, showing considerable variation in the methods' ability to explain the observed data. We also see differences in the experts' interpretation of similar levels of explanatory power. While the expert using PLS and PLSc interprets the values of 0.3 and higher as "satisfactory" considering prior research on related models, the CB-SEM expert takes similar values as "high" and "indicating strong effect sizes," referencing Cohen (1988). On the contrary, the experts using GSCA and $GSCA_M$ do not comment on the values. It is also worth noting that the latter two experts refer to explanatory power as "fit," a denotation grounded in econometric tradition but which contradicts the fit concept as understood in the classic psychometrics-based SEM literature (e.g., Henseler & Sarstedt, 2013).

4 | DISCUSSION OF RESULTS

Our consolidation of the experts' reports highlights results differences that make it very challenging for innovation management researchers and practitioners to draw a uniform picture from the analyses. Specifically, we observe divergent results regarding final model settings, the observed significances, and the relative importance of certain constructs for driving organizational innovation and performance.

On the one hand, these divergences may come as a surprise since one might expect the methods to perform similarly because they tap the same real-world phenomena. On the other hand, some differences can be expected as the methods rely on specific assumptions regarding the nature of the constructs (Sarstedt et al., 2016). While factor-based methods conceive a construct as an abstract entity that can in principle fully be captured by the covariation in the associated indicators, component-based methods seek to approximate constructs through linear combinations of indicator variables (Sarstedt et al., 2016). Methods also differ in terms of their optimization routines in that they either seek to minimize the divergence between sample-implied and model-implied covariances (Bollen, 1989, chap. 1; Diamantopoulos & Siguaw, 2000) or maximize explained variance (Cho et al., 2022). Researchers' preference for one method over the other, therefore, necessarily comes with assumptions about unknown entities in a model and the parameter estimation (Rigdon et al., 2017), giving rise to *methodological uncertainty*.

In addition, every method requires researchers to make specific choices in the model estimation process. For example, component-based SEM methods rely on bootstrapping for inference testing, which requires researchers to decide on the number of bootstrap samples and the confidence interval type, both of which have been shown to impact the model estimates (Aguirre-Urreta & Rönkkö, 2018; Streukens & Leroi-Werelds, 2016). Therefore, it is crucial to clearly document all parameter settings, which the experts in our analysis did to varying degrees. Specifically, the expert using PLS and PLSc estimated the model using the graphical user interface-based software SmartPLS (Ringle et al., 2024), reporting the version number and the use of the fixed seed option along with the number of bootstrap samples, thereby safeguarding reproducibility. The software used for GSCA and $GSCA_M$ estimations (GSCA Pro; Hwang et al., 2023) does not allow fixing the seed, which means that the bootstrapping routine produces slightly different results every time it is run. Finally, the expert using CB-SEM used syntax-based software, which generally facilitates the results' reproducibility, and made the assumptions underlying the model estimation transparent (i.e., use of robust maximum likelihood estimation due to ordinal data).[8]

A second source of uncertainty emerges from the analytical decisions that researchers make along the model estimation process (*model estimation uncertainty*). For example, we observed considerable divergence in how experts assessed the structural model. While some focused exclusively on model fit, others emphasized predictive power, or both. These divergences may partially be grounded in different traditions in applying the methods, particularly regarding establishing model fit versus predictive power (Evermann & Tate, 2016), but they clearly reflect different foci when running the analyses, which may not mirror current research on the methods. For example, different from the way the methods are portrayed in standard textbooks, the use of PLS generally allows for the use of model fit measures (Schuberth et al., 2023), just like prediction can be carried out in a factor-based SEM framework (de Rooij et al., 2022)—albeit with certain limitations (Hair et al., 2022, chap. 6).

Many of the analytical decisions that researchers make interact with each other such as a research project's general focus and the process of model evaluation. Specifically, researchers can assume different theoretical lenses in their projects by relying on theories designed solely for explanation, solely for prediction, and those encompassing both explanation and prediction (Gregor, 2006). The evaluation of the resulting models needs to consider

[8]Note that syntax-based estimation may also be subject to variability due to, for example, random algorithm initializations.

these different lenses by distinguishing between explanatory modeling, which emphasizes in-sample evaluations (e.g., explanatory power and model fit), and predictive modeling, which focuses on out-of-sample evaluations (e.g., predictive power). The principal distinction between these modeling paradigms lies in the fact that explanatory modeling facilitates ex-post inferences about how well a model explains or fits observed data, while predictive modeling supports ex-ante inferences about how well a model predicts or generalizes to hitherto unobserved data (Shmueli, 2010; Shmueli & Koppius, 2011). Distinguishing these perspectives is crucial, because a model that performs well in terms of explanation does not necessarily have high predictive power, and vice versa (Sarstedt & Danks, 2022). These different foci and the ensuing analysis steps may entail model modifications that result in different model configurations (e.g., to ensure improved predictive capabilities; Liengaard et al., 2021; Sharma et al., 2023).

Model modifications may not only be motivated on the grounds of statistical concerns, as was the case in the context of this article, but researchers may also explicitly hypothesize alternative models prior to the data analysis. Alternative models typically emerge when considering theories in new contexts with unique variables and effects, or when researchers build conceptual bridges across related streams of inquiry to provide a holistic understanding of the phenomenon (Burnham & Anderson, 2002; Sharma et al., 2019). Researchers then try to identify the model that best approximates the data generation process underlying the phenomenon under study. Such multimodel inference recognizes the practical reality that (1) researchers often face difficult choices among multiple competing hypotheses rather than just two mutually exclusive possibilities, and (2) the choice among the competitors is often not clear-cut so that researchers have an evidence-based reluctance to discard all but one (Rigdon et al., 2023). Unfortunately, such decision processes are rarely made transparent (John et al., 2012), but they should be.

Another source of uncertainty lies in the researchers' interpretation of the estimates, the context in which they are embedded (e.g., their relation to other variables or theoretical constructs), and common thresholds (*interpretational uncertainty*). For most individual paths in our example, all methods produce rather wide confidence intervals that overlap considerably. For example, the path coefficient estimates between acquisition and management innovation are almost equivalent across all methods but their confidence intervals' lower bounds vary around zero, entailing different interpretations concerning the path's significance. Furthermore, while the two path coefficient estimates are almost equivalent, either process or management innovation has a greater impact on organizational performance, depending on the method under consideration. These two examples demonstrate that accounting for variation can safeguard against overemphasizing small differences between point estimates or small deviations from thresholds. This issue has also been discussed in the context of model validation where researchers have shown that inferential tests produce considerably lower false positive rates when the estimate is equal to the assumed threshold (e.g., Franke & Sarstedt, 2019). In another stream of research, methodologists have called for complementing traditional significance testing with equivalence testing (Lakens et al., 2018) or Bayesian analyses (Wagenmakers, 2007). Bayesian analyses, for example, can estimate posterior probabilities and accumulate evidence in favor of or against the null hypothesis.

To summarize, the results presented here showcase a combination of three major sources of variability (Figure 5): methodological uncertainty, model estimation uncertainty, and interpretational uncertainty. In the following, we offer recommendations on how to address these uncertainty types in research projects.

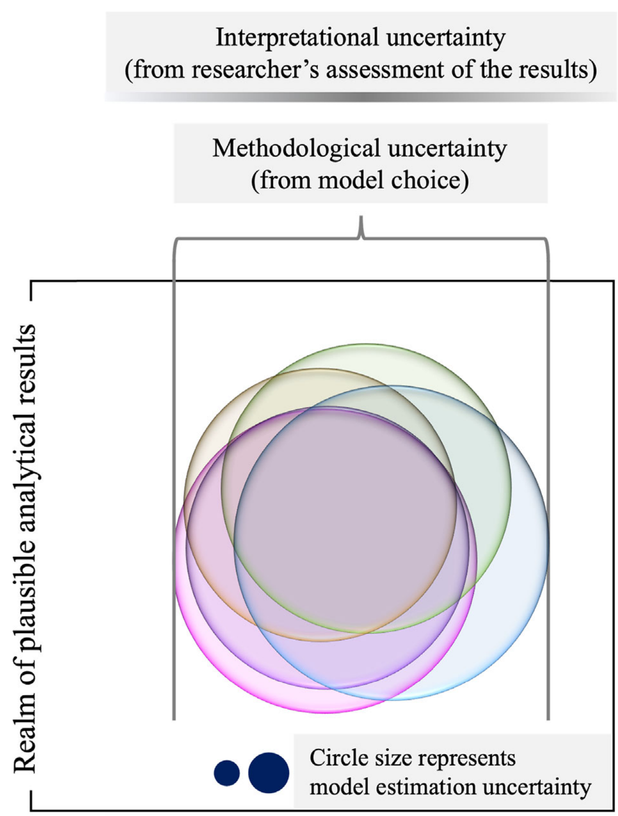

FIGURE 5 Schematic representation of three kinds of uncertainty.

5 | RECOMMENDATIONS

Uncertainty is inherent in the research process and cannot be avoided. This uncertainty goes well beyond random sampling error, as expressed in statistical standard errors, but has its origins in other elements of the research process (Rigdon & Sarstedt, 2022)—like those highlighted in our consolidation of the experts' reports. It comes with researchers' choice of a method to analyze the data and the numerous decisions they make before reporting the results and interpreting them. As Rigdon et al. (2020, p. 329) note, "uncertainty cannot be less than the standard error alone, unless researchers have ignored some phenomenon like finite population in computing their standard error." To grasp its potential impact, Rigdon et al. (2020) call for making all factors that may have contributed to this uncertainty transparent. In the context of our analysis, this would mean documenting choices made during the data analysis. Table 1 situates these choices in the analytical workflow and offers recommendations regarding their reporting. While the choices pertain to applications of SEM, many of them are broadly applicable to other methodological domains such as standard regression analyses. For example, rather than focusing on a single model while ignoring all the evidence that favored alternative models, researchers should assume different perspectives that may give rise to other configurations (Nuzzo, 2015), independent of the method used.

Making analytical choices transparent is important because they open a multitude of potential analytical workflows and alternative results (Wagenmakers et al., 2021). Choices such as whether to assess a model's predictive power as well as whether to omit or add paths open a garden of forking paths in that each analytical decision cuts off other possible routines and results (Gelman & Loken, 2014). A potential approach to exploit the analytical flexibility is a multiverse analysis where researchers document the results from different analytical workflows (Steegen et al., 2016). Researchers engaging in multiverse analyses specifically review possible analytical choices *before* the data analysis (e.g., which estimators or bootstrapping procedures to use) and identify reasonable choices for each analytical step. Instead of focusing on one workflow, a multiverse analysis applies *all* reasonable choices in alternative workflows, and

TABLE 1 Possible outline for capturing data analytical decisions in SEM and researcher's degrees of freedom.

Decision domains in the analytical workflow	Recommendations	Recommended reading
Method choice	Explicitly formulate and substantiate how the assumption regarding the constructs' nature (factors vs. components) and the goal of the analysis (explanation vs. prediction) align with method choice.	Cho et al. (2022); Gregor (2006); Hair and Sarstedt (2019); Rigdon et al. (2017)
Algorithmic implementation and settings	Document the concrete estimator and input format transparently. Document all algorithmic settings, including bootstrapping and related methods. Safeguard reproducibility, for example, by providing code or syntax, documenting software versions, etc.	CB-SEM: Diamantopoulos and Siguaw (2006) GSCA: Hwang and Takane (2004) $GSCA_M$: Hwang et al. (2017) PLS: Hair et al. (2021) PLSc: Henseler (2021)
Measurement model evaluation	Document the criteria applied, along with threshold values assumed in the analysis. Make any modification transparent.	
Structural model evaluation	Ensure that the model evaluation aligns with the focus of the analysis (explanation, prediction, or both). Document the selection of metrics and assumed threshold values. Document all modifications.	
Alternative models and methods	Identify alternative models and report their results. Make the genesis of the final model transparent, starting with the initial conceptual model. Document the model comparison process regarding the choice of metrics and theoretical values for each model.	Burnham and Anderson (2002); Liengaard et al. (2021); Sharma et al. (2023); Sharma et al. (2019)
Results interpretation	Report and interpret point estimates and their variation. Identify which results are unambiguous and outline inconclusive results.	Wagenmakers et al. (2021)

combines their results. Of course, multiverse analyses can quickly become very complex, rendering the documentation of all combinations and ensuing implications hardly feasible. Researchers should therefore focus on documenting the implications of major analytical choices. Conducting multiverse analyses requires researchers to embrace potentially inconclusive results if confidence intervals are wide or Bayes factors are uninformative. Such a step also implies abandoning the dichotomy of statistically significant versus not significant and, instead, interpreting the *p*-value as a continuous metric (McShane et al., 2023). As standard errors understate the overall uncertainty of results (Rigdon, 2023), researchers should demonstrate modesty and not oversell their results until they account for all material components of uncertainty (Rigdon et al., 2020).

Importantly, researchers should clearly label uncertain or inconclusive results to stimulate further research. Such practice also acknowledges that a single research article can never explain a phenomenon in full. It is a building block for accumulative science in which different research teams indirectly collaborate by sharing ideas, results, materials, and data through their publication activities.

6 | LIMITATIONS AND FURTHER RESEARCH

Our study showcases similarities and differences between the workflows and results that multiple experts employ when analyzing the same model using the same data. As with any research endeavor, we made several decisions during the study setup, analysis, and reporting that constitute limitations to our design.

We chose a specific workflow that applies one model and reviews expert analyses from five SEM methods. While this approach allowed us to showcase the rationales for different workflows in greater detail compared with more extensive crowdsourcing projects that primarily concern themselves with the convergence of statistical results (e.g., Silberzahn et al., 2018), it also restricts the garden of forking paths to a small subset of possible analytical decisions. Specifically, by using a prespecified model versus allowing other initial model specifications (e.g., using four dimensions of absorptive capacity vs. specifying a higher-order construct; see Zahra & George, 2002) and by providing a ready-to-use dataset, we constrained crucial theoretical and pre-analytical stages. These steps were necessary to ensure the comparability of the statistical results and workflows, but nullify other uncertainty components such as the measurement instruments and data (Rigdon & Sarstedt, 2022). It would therefore be worthwhile to extend the perspective by offering experts with more degrees of freedom, for example, by asking them to collect their own data. Relatedly, using alternative methods such as factor score regression (Dröge et al., 2000; Skrondal & Laake, 2001) or sum score regression (Duysters & Lokshin, 2011; Hair et al., 2024) may further increase the result's variability.

While making the decisions in the research process transparent and acknowledging alternative outcomes is important for managing components of uncertainty, only quantifying and managing uncertainty proactively will improve research practice over time. To do so, innovation researchers, and social science researchers in general, can draw on a rich methodological arsenal developed in metrology, which is the measurement science in physics, engineering, and legal forensics (JCGM, 2012). Researchers in these fields construct so-called uncertainty budgets that collect the available information on all material components of uncertainty and quantify their impact on the statistical estimators. These quantifications may result from replication studies, but may likewise be grounded in individual experience or scientific judgment (Bell, 1999). Rigdon et al. (2020) showcase the uncertainty quantification for a simple behavioral research experiment, but when it comes to analyses like those presented in this article, this process requires considerably more effort. For example, to limit uncertainty that comes with concept definitions and operationalizations, research fields may establish standard measures of concepts such as innovation orientation. Corresponding calls are not new (Elson et al., 2023; Rossiter, 2017), but they have not been echoed on a broader basis. There is no doubt that quantifying all material components of uncertainty requires a massive infrastructure that comes with substantial investments of time, money, and expertise. Science progresses one step at a time—the initial step is to acknowledge uncertainty components like those in our analyses and identify new ones.

7 | CONCLUSION

A single research article often only reports one out of a multitude of possible analytical workflows. We invited leading experts in the field of SEM—a prominent method in innovation research—to estimate a prespecified model on absorptive capacity's impact on organizational innovation and performance using the same data, but different estimators. Our results show that even when severely limiting the researchers' degrees of freedom in the application of the methods, the outcomes of the analyses are highly variable. Our findings suggest that researchers should acknowledge the uncertainty that comes with any statistical analysis and account for alternative analytical workflows to gradually improve research practice over time.

ACKNOWLEDGMENTS

The authors would like to thank Dr Murad Ali (Northumbria University Newcastle) for allowing the reuse and publication of the data. Open Access funding enabled and organized by Projekt DEAL.

FUNDING INFORMATION

The authors declare that no specific funding was received in support the research.

CONFLICT OF INTEREST STATEMENT

This research uses the statistical software SmartPLS (https://www.smartpls.com). Christian M. Ringle acknowledges a financial interest in SmartPLS.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available on: https://osf.io/29spz/

ETHICS STATEMENT

The authors have read and agreed to the Committee on Publication Ethics (COPE) international standards for authors.

ORCID

Marko Sarstedt https://orcid.org/0000-0002-5424-4268
Susanne J. Adler https://orcid.org/0000-0002-3211-6871
Christian M. Ringle https://orcid.org/0000-0002-7027-8804
Gyeongcheol Cho https://orcid.org/0000-0002-9237-0388
Adamantios Diamantopoulos https://orcid.org/0000-0001-5537-717X
Heungsun Hwang https://orcid.org/0000-0002-5057-7479
Benjamin D. Liengaard https://orcid.org/0000-0002-1657-3108

REFERENCES

Aguirre-Urreta, Miguel I., and Mikko Rönkkö. 2018. "Statistical Inference with PLSc Using Bootstrap Confidence Intervals." *MIS Quarterly* 42(3): 1001–20. https://doi.org/10.25300/misq/2018/13587.

Alesanco-Llorente, María, Eva Reinares-Lara, Jorge Pelegrín-Borondo, and Cristina Olarte-Pascual. 2023. "Mobile-Assisted Showrooming Behavior and the (R)evolution of Retail: The Moderating Effect of Gender on the Adoption of Mobile Augmented Reality." *Technological Forecasting and Social Change* 191: 122514. https://doi.org/10.1016/j.techfore.2023.122514.

Ali, Murad, Konan Anderson Seny Kan, and Marko Sarstedt. 2016. "Direct and Configurational Paths of Absorptive Capacity and Organizational Innovation to Successful Organizational Performance." *Journal of Business Research* 69(11): 5317–23. https://doi.org/10.1016/j.jbusres.2016.04.131.

Anderson, James C., and David W. Gerbing. 1988. "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach." *Psychological Bulletin* 103(3): 411–423. https://doi.org/10.1037/0033-2909.103.3.411.

Bagozzi, Richard P., and Youjae Yi. 1988. "On the Evaluation of Structural Equation Models." *Journal of the Academy of Marketing Science* 16(1): 74–94. https://doi.org/10.1007/BF02723327.

Bell, Stephanie. 1999. *Good Practice Guide #12: A Beginner's Guide to Uncertainty of Measurement*. Teddington, UK: National Physical Laboratory.

Berndt, Ana Clara, Giancarlo Gomes, and Felipe Mendes Borini. 2023. "Exploring the Antecedents of Frugal Innovation and Operational Performance: The Role of Organizational Learning Capability and Entrepreneurial Orientation." *European Journal of Innovation Management*. https://doi.org/10.1108/EJIM-06-2022-0320.

Bollen, Kenneth A. 1989. *Structural Equations with Latent Variables*. New York, NY: Wiley.

Botvinik-Nezer, Rotem, Felix Holzmeister, Colin F. Camerer, Anna Dreber, Juergen Huber, Magnus Johannesson, M. Kirchler, et al. 2020. "Variability in the Analysis of a Single Neuroimaging Dataset by Many Teams." *Nature* 582(7810): 84–88. https://doi.org/10.1038/s41586-020-2314-9.

Breznau, Nate, Eike M. Rinke, Alexander Wuttke, Hung H. V. Nguyen, Muna Adem, Jule Adriaans, A. Alvarez-Benjumea, et al. 2022. "Observing Many Researchers Using the Same Data and Hypothesis Reveals a Hidden Universe of Uncertainty." *Proceedings of the National Academy of Sciences* 119(44): e2203150119. https://doi.org/10.1073/pnas.2203150119.

Buchanan, John T., Erez J. Henig, and Mordecai I. Henig. 1998. "Objectivity and Subjectivity in the Decision Making Process." *Annals of Operations Research* 80: 333–345. https://doi.org/10.1023/A:1018980318183.

Burnham, Kenneth P., and David R. Anderson. 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. New York, NY: Springer.

Cautela, Cabirio, Michele Simoni, and Peter Moran. 2022. "Microfoundations of Dynamic Design Capabilities: An Empirical Analysis of "Excellent" Italian Design Firms." *Journal of Product Innovation Management* 39(1): 3–23. https://doi.org/10.1111/jpim.12592.

Chin, Wynne W. 1998. "The Partial Least Squares Approach to Structural Equation Modeling." In *Modern Methods for Business Research*, edited by George A. Marcoulides, 295–358. Mahwah, NJ: Lawrence Erlbaum.

Cho, Gyeongcheol, Heungsun Hwang, Marko Sarstedt, and Christian M. Ringle. 2020. "Cutoff Criteria for Overall Model Fit Indexes in Generalized Structured Component Analysis." *Journal of Marketing Analytics* 8(4): 189–202. https://doi.org/10.1057/s41270-020-00089-1.

Cho, Gyeongcheol, Jonathan Lee, Heungsun Hwang, Marko Sarstedt, and Christian M. Ringle. 2023. "A Comparative Study of the Predictive Power of Component-Based Approaches to Structural Equation Modeling." *European Journal of Marketing* 57(6): 1641–61. https://doi.org/10.1108/EJM-07-2020-0542.

Cho, Gyeongcheol, Marko Sarstedt, and Heungsun Hwang. 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* 73(2): 220–251. https://doi.org/10.1111/bmsp.12255.

Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Mahwah, NJ: Lawrence Erlbaum.

Cohen, Wesley M., and Daniel A. Levinthal. 1990. "Absorptive Capacity: A New Perspective on Learning and Innovation." *Administrative Science Quarterly* 35(1): 128–152. https://doi.org/10.2307/2393553.

Cronbach, Lee J. 1951. "Coefficient Alpha and the Internal Structure of Tests." *Psychometrika* 16(3): 297–334. https://doi.org/10.1007/BF02310555.

de Rooij, Mark, Julian D. Karch, Marjolein Fokkema, Zsuzsa Bakk, Bunga Citra Pratiwi, and Henk Kelderman. 2022. "SEM-Based Out-of-Sample Predictions." *Structural Equation Modeling: A Multidisciplinary Journal* 30(1): 132–148. https://doi.org/10.1080/10705511.2022.2061494.

Diamantopoulos, Adamantios, and Judy A. Siguaw. 2000. *Introducing LISREL*. Thousand Oaks, CA: Sage.

Diamantopoulos, Adamantios, and Judy A. Siguaw. 2006. "Formative Versus Reflective Indicators in Organizational Measure Development: A Comparison and Empirical Illustration." *British Journal of Management* 17(4): 263–282. https://doi.org/10.1111/j.1467-8551.2006.00500.x.

Dijkstra, Theo K. 2010. "Latent Variables and Indices: Herman Wold's Basic Design and Partial Least Squares." In *Handbook of Partial Least Squares: Concepts, Methods and Applications (Springer Handbooks of Computational Statistics Series, Vol. II)*, edited by Vincenzo Esposito Vinzi, Wynne W. Chin, Jörg Henseler, and Huiwen Wang, 23–46. Heidelberg, Dordrecht, London, New York: Springer. https://doi.org/10.1007/978-3-540-32827-8_2.

Dijkstra, Theo K., and Jörg Henseler. 2015. "Consistent Partial Least Squares Path Modeling." *MIS Quarterly* 39(2): 297–316. https://doi.org/10.25300/MISQ/2015/39.2.02.

Dijkstra, Theo K., and Karin Schermelleh-Engel. 2014. "Consistent Partial Least Squares for Nonlinear Structural Equation Models." *Psychometrika* 79(4): 585–604. https://doi.org/10.1007/s11336-013-9370-0.

Dröge, Cornelia, Jayanth Jayaram, and Shawnee K. Vickery. 2000. "The Ability to Minimize the Timing of New Product Development and Introduction: An Examination of Antecedent Factors in the North American Automobile Supplier Industry." *Journal of Product Innovation Management* 17(1): 24–40. https://doi.org/10.1016/S0737-6782(99)00009-0.

Duysters, Geert, and Boris Lokshin. 2011. "Determinants of Alliance Portfolio Complexity and its Effect on Innovative Performance of Companies." *Journal of Product Innovation Management* 28(4): 570–585. https://doi.org/10.1111/j.1540-5885.2011.00824.x.

Elson, Malte, Ian Hussey, Taym Alsalti, and Ruben C. Arslan. 2023. "Psychological Measures Aren't Toothbrushes." *Communications Psychology* 1: 25. https://doi.org/10.1038/s44271-023-00026-9.

Evermann, Joerg, and Mary Tate. 2016. "Assessing the Predictive Performance of Structural Equation Model Estimators." *Journal of Business Research* 69(10): 4565–82. https://doi.org/10.1016/j.jbusres.2016.03.050.

Fiss, Peer C. 2011. "Building Better Causal Theories: A Fuzzy Set Approach to Typologies in Organization Research." *Academy of Management Journal* 54(2): 393–420. https://doi.org/10.5465/amj.2011.60263120.

Franke, George, and Marko Sarstedt. 2019. "Heuristics Versus Statistics in Discriminant Validity Testing: A Comparison of Four Procedures." *Internet Research* 29(3): 430–447. https://doi.org/10.1108/IntR-12-2017-0515.

Gelman, Andrew, and Eric Loken. 2014. "The Statistical Crisis in Science." *American Scientist* 102(6): 460–65. https://doi.org/10.1511/2014.111.460.

Gregor, Shirley. 2006. "The Nature of Theory in Information Systems." *MIS Quarterly* 30(3): 611–642. https://doi.org/10.2307/25148742.

Hair, Joseph F., Matt C. Howard, and Christian Nitzl. 2020. "Assessing Measurement Model Quality in PLS-SEM Using Confirmatory Composite Analysis." *Journal of Business Research* 109: 101–110.

Hair, Joseph F., and Marko Sarstedt. 2019. "Factors Versus Composites: Guidelines for Choosing the Right Structural Equation Modeling Method." *Project Management Journal* 50(6): 619–624. https://doi.org/10.1177/8756972819882132.

Hair, Joseph F., Pratyush N. Sharma, M. Sarstedt, Christian M. Ringle, and Benjamin D. Liengaard. 2024. "The Shortcomings of Equal Weights Estimation and the Composite Equivalence Index in PLS-SEM." *European Journal of Marketing* 58(13): 30–55. https://doi.org/10.1108/EJM-04-2023-0307.

Hair, Joseph F., G. Tomas, M. Hult, Christian M. Ringle, M. Sarstedt, Nicholas P. Danks, and Soumya Ray. 2021. *Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R*. Cham: Springer.

Hair, Joseph F., G. Tomas, M. Hult, Christian M. Ringle, and Marko Sarstedt. 2022. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*, Vol 3. Thousand Oaks, CA: Sage.

Hair, Joseph F., G. Tomas, M. Hult, Christian M. Ringle, Marko Sarstedt, and Kai O. Thiele. 2017. "Mirror, Mirror on the Wall: A Comparative Evaluation of Composite-Based Structural Equation Modeling Methods." *Journal of the Academy of Marketing Science* 45(5): 616–632. https://doi.org/10.1007/s11747-017-0517-x.

Henseler, Jörg. 2021. *Composite-Based Structural Equation Modeling Analyzing Latent and Emergent Variables*. New York, NY: Guilford Press.

Henseler, Jörg, Christian M. Ringle, and Marko Sarstedt. 2015. "A New Criterion for Assessing Discriminant Validity in Variance-Based Structural Equation Modeling." *Journal of the Academy of Marketing Science* 43(1): 115–135. https://doi.org/10.1007/s11747-014-0403-8.

Henseler, Jörg, and Marko Sarstedt. 2013. "Goodness-of-Fit Indices for Partial Least Squares Path Modeling." *Computational Statistics* 28(2): 565–580. https://doi.org/10.1007/s00180-012-0317-1.

Hwang, Heungsun, Gyeongcheol Cho, and Hosung Choo. 2023. "GSCA Pro—Free Stand-Alone Software for Structural Equation Modeling." *Structural Equation Modeling*. https://doi.org/10.1080/10705511.2023.2272294.

Hwang, Heungsun, Gyeongcheol Cho, Kwanghee Jung, Carl F. Falk, Jessica Kay Flake, Min Jin Jin, and Seung Hwan Lee. 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, Heungsun, Marko Sarstedt, Jun Hwa Cheah, and Christian M. Ringle. 2020. "A Concept Analysis of Methodological Research on Composite-Based Structural Equation Modeling: Bridging PLSPM and GSCA." *Behaviormetrika* 47(1): 219–241. https://doi.org/10.1007/s41237-019-00085-5.

Hwang, Heungsun, and Yoshio Takane. 2004. "Generalized Structured Component Analysis." *Psychometrika* 69(1): 81–99. https://doi.org/10.1007/bf02295841.

Hwang, Heungsun, and Yoshio Takane. 2014. *Generalized Structured Component Analysis*. New York, NY: Chapman and Hall/CRC.

Hwang, Heungsun, Yoshio Takane, and Kwanghee Jung. 2017. "Generalized Structured Component Analysis with Uniqueness Terms for Accommodating Measurement Error." *Frontiers in Psychology* 8: 2137. https://doi.org/10.3389/fpsyg.2017.02137.

JCGM, ed. 2012. *International Vocabulary of Metrology—Basic and General Concepts and Associated Terms (VIM)*, 3rd ed. Sèvres: JCGM.

Jean, Ruey-Jer "Bryan", Rudolf R. Sinkovics, and Thomas P. Hiebaum. 2014. "The Effects of Supplier Involvement and Knowledge Protection on Product Innovation in Customer-Supplier Relationships: A Study of Global Automotive Suppliers in China." *Journal of Product Innovation Management* 31(1): 98–113. https://doi.org/10.1111/jpim.12082.

John, Leslie K., George Loewenstein, and Drazen Prelec. 2012. "Measuring the Prevalence of Questionable Research Practices with Incentives for Truth Telling." *Psychological Science* 23(5): 524–532. https://doi.org/10.1177/0956797611430953.

Jöreskog, Karl G. 1978. "Structural Analysis of Covariance and Correlation Matrices." *Psychometrika* 43(4): 443–477. https://doi.org/10.1007/bf02293808.

Jöreskog, Karl G., Ulf H. Olsson, and Fan Y. Wallentin. 2016. "Confirmatory Factor Analysis (CFA)." In *Multivariate Analysis with LISREL*, edited by Karl G. Jöreskog and Ulf H. Olsson, 283–339. Cham: Springer. https://doi.org/10.1007/978-3-319-33153-9_7.

Kemper, Jan, Oliver Schilke, and Malte Brettel. 2013. "Social Capital as a Microlevel Origin of Organizational Capabilities." *Journal of Product Innovation Management* 30(3): 589–603. https://doi.org/10.1111/jpim.12004.

Lakens, Daniël, Anne M. Scheel, and Peder M. Isager. 2018. "Equivalence Testing for Psychological Research: A Tutorial." *Advances in Methods and Practices in Psychological Science* 1(2): 259–269. https://doi.org/10.1177/2515245918770963.

Liengaard, Benjamin Dybro, Pratyush Nidhi Sharma, G. Tomas, M. Hult, Morten Berg Jensen, Marko Sarstedt, Joseph F. Hair, and Christian M. Ringle. 2021. "Prediction: Coveted, Yet Forsaken? Introducing a Cross-Validated Predictive Ability Test in Partial Least Squares Path Modeling." *Decision Sciences* 52(2): 362–392. https://doi.org/10.1111/deci.12445.

Lohmöller, Jan-Bernd. 1989. *Latent Variable Path Modeling With Partial Least Squares*. Heidelberg: Springer.

McShane, Blakeley B., Eric T. Bradlow, John G. Lynch, and Robert J. Meyer. 2023. ""Statistical Significance" and Statistical Reporting: Moving beyond Binary." *Journal of Marketing*. https://doi.org/10.1177/00222429231216910.

Nuzzo, Regina. 2015. "How Scientists Fool Themselves– and How They Can Stop." *Nature* 526(7572): 182–85. https://doi.org/10.1038/526182a.

Paxton, Pamela, Patrick J. Curran, Kenneth A. Bollen, Jim Kirby, and Feinian Chen. 2001. "Monte Carlo Experiments: Design and Implementation." *Structural Equation Modeling: A Multidisciplinary Journal* 8(2): 287–312. https://doi.org/10.1207/s15328007sem0802_7.

Pemartín, María, Ana I. Rodríguez-Escudero, and José Luís Munuera-Alemán. 2018. "Effects of Collaborative Communication on NPD Collaboration Results: Two Routes of Influence." *Journal of Product Innovation Management* 35(2): 184–208. https://doi.org/10.1111/jpim.12375.

Reinartz, Werner J., Michael Haenlein, and Jörg Henseler. 2009. "An Empirical Comparison of the Efficacy of Covariance-Based and Variance-Based SEM." *International Journal of Research in Marketing* 26(4): 332–344. https://doi.org/10.1016/j.ijresmar.2009.08.001.

Rigdon, Edward E. 1998. "Structural Equation Modeling." In *Modern Methods for Business Research*, edited by George A. Marcoulides, 251–294. Mahwah, NJ: Lawrence Erlbaum.

Rigdon, Edward E. 2023. "How Improper Dichotomization and the Misrepresentation of Uncertainty Undermine Social Science Research." *Journal of Business Research* 165: 114086. https://doi.org/10.1016/j.jbusres.2023.114086.

Rigdon, Edward E., and Marko Sarstedt. 2022. "Accounting for Uncertainty in the Measurement of Unobservable Marketing Phenomena." In *Review of Marketing Research*, edited by Hans Baumgartner and Bert Wejters, 53–73. Emerald: Bingley.

Rigdon, Edward E., Marko Sarstedt, and Jan-Michael Becker. 2020. "Quantify Uncertainty in Behavioral Research." *Nature Human Behaviour* 4: 329–331. https://doi.org/10.1038/s41562-019-0806-0.

Rigdon, Edward E., Marko Sarstedt, and Christian M. Ringle. 2017. "On Comparing Results from CB-SEM and PLS-SEM: Five Perspectives and Five Recommendations." *Marketing ZFP—Journal of Research and Management* 39(3): 4–16. https://doi.org/10.15358/0344-1369-2017-3-4.

Rigdon, Edward, Marko Sarstedt, and Ovidiu-Ioan Moisescu. 2023. "Quantifying Model Selection Uncertainty Via Bootstrapping and Akaike Weights." *International Journal of Consumer Studies* 47(4): 1596–1608. https://doi.org/10.1111/ijcs.12906.

Ringle, Christian M., Sven Wende, and Jan-Michael Becker. 2024. *SmartPLS 4*. Bönningstedt: SmartPLS GmbH. https://www.smartpls.com.

Rossiter, John R. 2017. "Optimal Standard Measures for Marketing." *Journal of Marketing Management* 33(5-6): 313–326. https://doi.org/10.1080/0267257X.2017.1293710.

Sarstedt, Marko, and Nicholas P. Danks. 2022. "Prediction in HRM Research—A Gap between Rhetoric and Reality." *Human Resource Management Journal* 32(2): 485–513. https://doi.org/10.1111/1748-8583.12400.

Sarstedt, Marko, Joseph F. Hair, Christian M. Ringle, Kai O. Thiele, and Siegfried P. Gudergan. 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, Florian, Jörg Henseler, and Theo K. Dijkstra. 2018. "Confirmatory Composite Analysis." *Frontiers in Psychology* 9: 2541. https://doi.org/10.3389/fpsyg.2018.02541.

Schuberth, Florian, Manuel E. Rademaker, and Jörg Henseler. 2023. "Assessing the Overall Fit of Composite Models

Estimated by Partial Least Squares Path Modeling." *European Journal of Marketing* 57(6): 1678–1702. https://doi.org/10.1108/EJM-08-2020-0586.

Schweinsberg, Martin, Michael Feldman, Nicola Staub, Olmo R. van den Akker, Robbie C. M. van Aert, Marcel A. L. M. van Assen, Yang Liu, et al. 2021. "Same Data, Different Conclusions: Radical Dispersion in Empirical Results when Independent Analysts Operationalize and Test the Same Hypothesis." *Organizational Behavior and Human Decision Processes* 165: 228–249. https://doi.org/10.1016/j.obhdp.2021.02.003.

Sharma, Pratyush N., Benjamin D. Liengaard, Joseph F. Hair, Marko Sarstedt, and Christian M. Ringle. 2023. "Predictive Model Assessment and Selection in Composite-Based Modeling Using PLS-SEM: Extensions and Guidelines for Using CVPAT." *European Journal of Marketing* 57(6): 1662–77. https://doi.org/10.1108/ejm-08-2020-0636.

Sharma, Pratyush N., M. Sarstedt, Galit Shmueli, and Kai O. Thiele. 2019. "PLS-Based Model Selection: The Role of Alternative Explanations in IS Research." *Journal of the Association for Information Systems* 20(4): 346–397. https://doi.org/10.17005/1.jais.00538.

Shmueli, Galit. 2010. "To Explain or to Predict?" *Statistical Science* 25(3): 289–310. https://doi.org/10.1214/10-sts330.

Shmueli, Galit, and Otto R. Koppius. 2011. "Predictive Analytics in Information Systems Research." *MIS Quarterly* 35(3): 553–572. https://doi.org/10.2307/23042796.

Siguaw, Judy A., Penny M. Simpson, and Cathy A. Enz. 2006. "Conceptualizing Innovation Orientation: A Framework for Study and Integration of Innovation Research." *Journal of Product Innovation Management* 23(6): 556–574. https://doi.org/10.1111/j.1540-5885.2006.00224.x.

Silberzahn, R., E. L. Uhlmann, D. P. Martin, P. Anselmi, F. Aust, E. Awtrey, S. Bahnik, et al. 2018. "Many Analysts, One Data Set: Making Transparent How Variations in Analytic Choices Affect Results." *Advances in Methods and Practices in Psychological Science* 1(3): 337–356. https://doi.org/10.1177/2515245917747646.

Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn. 2011. "False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant." *Psychological Science* 22(11): 1359–66. https://doi.org/10.1177/0956797611417632.

Skrondal, Anders, and Petter Laake. 2001. "Regression Among Factor Scores." *Psychometrika* 66(4): 563–575. https://doi.org/10.1007/bf02296196.

Steegen, Sara, Francis Tuerlinckx, Andrew Gelman, and Wolf Vanpaemel. 2016. "Increasing Transparency Through a Multiverse Analysis." *Perspectives on Psychological Science* 11(5): 702–712. https://doi.org/10.1177/1745691616658637.

Stock, Ruth Maria, and Nicolas Andy Zacharias. 2011. "Patterns and Performance Outcomes of Innovation Orientation." *Journal of the Academy of Marketing Science* 39(6): 870–888. https://doi.org/10.1007/s11747-010-0225-2.

Streukens, Sandra, and Sara Leroi-Werelds. 2016. "Bootstrapping and PLS-SEM: A Step-by-Step Guide to Get More out of your Bootstrap Results." *European Management Journal* 34(6): 618–632. https://doi.org/10.1016/j.emj.2016.06.003.

Suder, Marcin, Joanna Duda, Rafał Kusa, and Alexandra Mora-Cruz. 2022. "At the Crossroad of Digital and Tourism Entrepreneurship: Mediating Effect of Digitalization in Hospitality Industry." *European Journal of Innovation Management*. https://doi.org/10.1108/EJIM-08-2022-0422.

Tenenhaus, Michel. 2008. "Component-Based Structural Equation Modelling." *Total Quality Management & Business Excellence* 19(7-8): 871–886. https://doi.org/10.1080/14783360802159543.

Wagenmakers, Eric-Jan. 2007. "A Practical Solution to the Pervasive Problems of P Values." *Psychonomic Bulletin & Review* 14(5): 779–804. https://doi.org/10.3758/bf03194105.

Wagenmakers, Eric-Jan, Alexandra Sarafoglou, Sil Aarts, Casper Albers, Johannes Algermissen, Štěpán Bahnik, N. von Dongan, et al. 2021. "Seven Steps toward More Transparency in Statistical Practice." *Nature Human Behaviour* 5(11): 1473–80. https://doi.org/10.1038/s41562-021-01211-8.

Wagenmakers, Eric-Jan, Alexandra Sarafoglou, and Balazs Aczel. 2022. "One Statistical Analysis Must Not Rule Them All." *Nature* 605: 423–25. https://doi.org/10.1038/d41586-022-01332-8.

Wiesböck, Florian, Thomas Hess, and Jelena Spanjol. 2020. "The Dual Role of IT Capabilities in the Development of Digital Products and Services." *Information & Management* 57(8): 103389. https://doi.org/10.1016/j.im.2020.103389.

Wold, Herman. 1982. "Soft Modeling: The Basic Design and Some Extensions." In *Systems Under Indirect Observations: Part II*, edited by Karl G. Jöreskog and Herman Wold, 1–54. Amsterdam: North-Holland.

Zahra, Shaker A., and Gerard George. 2002. "Absorptive Capacity: A Review, Reconceptualization, and Extension." *The Academy of Management Review* 27(2): 185–203. https://doi.org/10.2307/4134351.

AUTHOR BIOGRAPHIES

Marko Sarstedt (sarstedt@lmu.de) is a Chaired Professor of Marketing at the Ludwig-Maximilians-University Munich (Germany) and an adjunct research professor at Babeş-Bolyai-University Cluj-Napoca (Romania). His main research interest is the advancement of research methods to further the understanding of consumer behavior. His research has been published in Nature Human Behaviour, Journal of Marketing Research, Journal of the Academy of Marketing Science, Multivariate Behavioral Research, Organizational Research Methods, MIS Quarterly, and Psychometrika, among others. Since 2018, Marko has been named member of Clarivate Analytics' Highly Cited Researchers List, which includes the "world's most impactful scientific researchers."

Susanne J. Adler (adler@lmu.de) is a Research Associate at the Institute of Marketing at Ludwig-Maximilians-Universität in Munich, Germany. Her research focuses on how context affects consumer decision-making and meta-science. She has published on these topics in journals such as the Journal of Business Research and Psychology & Marketing.

Christian M. Ringle (c.ringle@tuhh.de) is a Chaired Professor and the Director of the Institute of Management and Decision Sciences at the Hamburg University of Technology (Germany) and an Adjunct Professor at the James Cook University (Australia). His research focuses on management and marketing topics, method development, business analytics, machine learning, and the application of business research methods to decision-making. Christian's contributions have been published in journals such as the Decision Sciences, European Journal of Marketing, International Journal of Research in Marketing, Information Systems Research, Journal of Business Research, Journal of Service Research, Journal of the Academy of Marketing Science, Long Range Planning, and MIS Quarterly. Since 2018, Christian has been included in the Clarivate Analytics' Highly Researchers list. He is a co-developer and cofounder of the statistical software SmartPLS.

Gyeongcheol Cho (cho.1240@osu.edu) is an Assistant Professor of Psychology at The Ohio State University. His research endeavors to advance quantitative methods for studying hypothetical relationships among behavioral, psychological, and biological variables, particularly through the measurement of theoretical constructs. His technical expertise includes factor/component analyses, structural equation modeling, and integrating these approaches with machine learning. He has published extensively in esteemed journals within the quantitative domain, such as Psychological Methods, Structural Equation Modeling: A Multidisciplinary Journal, Multivariate Behavioral Research, the British Journal of Mathematical and Statistical Psychology, and Psychometrika, among others.

Adamantios Diamantopoulos (adamantios.diamantopoulos@univie.ac.at) is a professorial research fellow at the University of Vienna, Austria, where he previously held the Chair of International Marketing (2004–2023). His main research interests are in international marketing and quantitative methods and his work has been published, among others, in the Journal of Marketing Research, Journal of International Business Studies, Journal of the Academy of Marketing Science, International Journal of Research in Marketing, Journal of Service Research, Journal of International Marketing, Journal of Retailing, MIS Quarterly, Organizational Research Methods, Psychological Methods, Information Systems Research, and British Journal of Management.

Heungsun Hwang (heungsun.hwang@mcgill.ca) is a Professor of Psychology at McGill University and the Director of the Psychological Data Analytics Laboratory. Before joining McGill, he was also an Assistant Professor of Marketing at HEC Montreal. He serves on the editorial boards of various journals, including Behaviormetrika, British Journal of Mathematical and Statistical Psychology, Ethics and Behavior, and Psychometrika. His research is devoted to the development and application of advanced quantitative analytics for the measurement and analysis of human characteristics, aspects, and processes.

Benjamin D. Liengaard (benlien@econ.au.dk) is an Associate Professor at the Department of Economics and Business Economics, at Aarhus University, Denmark. His research interest relates to partial least squares structural equation modeling, prediction-oriented model evaluation, the handling of endogeneity in regression models, and the use of textual analysis to quantify economic narratives. Benjamin has disseminated his work through journal articles, including the European Journal of Marketing, Decision Sciences, Psychology & Marketing, and the Journal of Applied Econometrics. He is an active reviewer for international journals.

How to cite this article: Sarstedt, Marko, Susanne J. Adler, Christian M. Ringle, Gyeongcheol Cho, Adamantios Diamantopoulos, Heungsun Hwang, and Benjamin D. Liengaard. 2024. "Same Model, Same Data, but Different Outcomes: Evaluating the Impact of Method Choices in Structural Equation Modeling." *Journal of Product Innovation Management* 1–17. https://doi.org/10.1111/jpim.12738