



Confirmatory composite analysis using partial least squares: setting the record straight

Florian Schubert¹

Received: 16 March 2020 / Accepted: 1 July 2020
© The Author(s) 2020

Abstract

Confirmatory composite analysis (CCA) is a subtype of structural equation modeling that assesses composite models. Composite models consist of a set of inter-related emergent variables, i.e., constructs which emerge as linear combinations of other variables. Only recently, Hair et al. (*J Bus Res* 109(1):101–110, 2020) proposed ‘confirmatory composite analysis’ as a method of confirming measurement quality (MCMQ) in partial least squares structural equation modeling. As a response to their study and to prevent researchers from confusing the two, this article explains what CCA and MCMQ are, what steps they entail and what differences they have. Moreover, to demonstrate their efficacy, a scenario analysis was conducted. The results of this analysis imply that to assess composite models, researchers should use CCA, and to assess reflective and causal–formative measurement models, researchers should apply structural equation modeling including confirmatory factor analysis instead of Hair et al.’s MCMQ. Finally, the article offers a set of corrections to the article of Hair et al. (2020) and stresses the importance of ensuring that the applied model assessment criteria are consistent with the specified model.

Keywords Confirmatory composite analysis · CCA · Composite model · Method of confirming measurement quality · Structural equation modeling · Latent variables · Emergent variables · Model fit assessment · PLS-SEM

Mathematics Subject Classification 62H99 · 62P25 · 91C99

This paper reuses material from Schubert et al.’s (2018) original publication on confirmatory composite analysis: <https://doi.org/10.3389/fpsyg.2018.02541>.

✉ Florian Schubert
f.schubert@utwente.nl

¹ Faculty of Engineering Technology, University of Twente, Drienerlolaan 5, 7522 NB Enschede, The Netherlands

1 Introduction

Marketing research aims at explaining the exchange between marketers and customers (Bagozzi 1975). Hence, Marketing comprises both a design component, how things *should* be done, and a behavioral component, what *do* actors in the marketplace do (Hunt 2010). In specific, Marketing research develops strategies such as marketing mix (McCarthy 1978) and market orientation (Kohli and Jaworski 1990), wherefore it can be regarded as a “science of the artificial” (Simon 1969). Moreover, it investigates the effects of these human-made concepts, so-called artifacts on, for example, the customer’s behavior. Hence, it can be regarded as a Behavioral science. As a consequence, it must combine the research paradigms of both Design and Behavioral research (Henseler 2017).

To meet the demands of both Design and Behavioral sciences in empirical research, structural equation modeling (SEM) is eminently suitable. It allows to model abstract concepts by a set of observable variables and to connect constructs, i.e., the statistical representation of the concepts, via a structural model (Bollen 1989). In doing so, complex relationships among the constructs and random measurement errors in the observable variables can be taken into account. Hence, SEM is a flexible approach that allows researchers to statistically model their theories. Moreover, researchers can use SEM to empirically falsify their theories which makes it a favorable tool in many disciplines including the field of Marketing research (Baumgartner and Homburg 1996; Steenkamp and Baumgartner 2000).

In SEM, three ways of modeling abstract concepts have been established: the reflective measurement model, the causal–formative measurement model, and the composite model (see e.g., Bollen and Bauldry 2011). Both, the reflective and causal–formative measurement model operationalize the concept under investigation as a latent variable, which has been proven to be particularly valuable for concepts of Behavioral sciences such as attitudes and personality traits. While in the reflective measurement model, it is assumed that the concept manifests itself in its measures, i.e., the measures share a common cause (Reichenbach 1956), in the causal–formative measurement model, the observable variables are assumed to cause the concept (Diamantopoulos et al. 2008). In contrast, the composite model can be used to operationalize abstract concepts that are conceived as constructions, i.e., as artifacts, which are objects designed to fulfill a certain purpose (Hevner et al. 2004; Henseler 2015, 2017; Schubert et al. 2018; Benitez et al. 2020).¹ In doing so, the concept is modeled as an emergent variable, i.e., instead of a latent variable, a linear combination of observable variables represents the concept in the statistical model (Cohen et al. 1990; Reise 1999). Hence, it is assumed that a combination of ingredients compose the artifact, which, in turn, is represented by an emergent variable in the model.

¹ Besides using the composite model to operationalize artifacts, it is acknowledged that both latent variables and emergent variables can be used to operationalize theoretical concepts of Behavioral sciences (Rigdon 2016; Rigdon et al. 2017, 2019; Rhemtulla et al. 2020). This does not limit the contributions of the study at hand.

Although emergent variables are less often encountered in empirical research than latent variables to model abstract concepts, they have been acknowledged and employed in various disciplines. For example concepts like marketing mix known from Marketing research (Fornell and Bookstein 1982) and concepts from Information Systems research such as IT infrastructure flexibility (Benitez et al. 2018) and IT integration capability (Braojos et al. 2020) have been modeled as emergent variable. Moreover, in Tourism and Hospitality research and Ecology, concepts such as tourist engagement (Rasoolimanesh et al. 2019) and plant community structure (Antoninka et al. 2011) were represented in the statistical model by an emergent variable. Similarly, in Evolution and Human Behavior research, the concept of environmental harshness was modeled as an emergent variable (Mell et al. 2018).

To broaden the accessibility of SEM to composite models, and therefore to open it to sciences of the artificial, SEM has been extended recently by confirmatory composite analysis (CCA; Henseler et al. 2014; Schubert et al. 2018). For a recent introduction of CCA to Business research, see Henseler and Schubert (forthcoming). CCA is a subtype of SEM, and thus follows the typical steps of SEM, namely, model specification, model identification, model estimation, and model assessment. To specify a model in CCA, emergent variables are employed and interrelated to represent the abstract concepts of interest. After the model has been specified and it has been ensured that the model is identified, i.e., the parameters can be uniquely from the variance–covariance matrix of the observable variables, the model parameters need to be estimated. In doing so, an estimator is employed that conforms the assumptions implied by the specified composite model. Finally, the estimated composite model is assessed globally and locally, i.e., the overall model fit and each emergent variable with its respective observable variables are examined. Consequently, CCA is similar to CFA, but instead of empirically falsifying reflective measurement models, it empirically falsifies composite models (Schubert et al. 2018).

For SEM various estimators have been proposed (Reinartz et al. 2009). Particularly, the partial least squares (PLS) estimator [as developed by Wold (1975)] has gained increasing attention over the last two decades (Henseler et al. 2009; Hair et al. 2012; Shiau et al. 2019). PLS is a two-step estimation approach, which, in the first step, obtains weights by an iterative algorithm to create linear combinations of observable variables that serve as proxies for the constructs in the structural model. Subsequently, in the second step, the path coefficients are estimated by ordinary least squares based on the correlation matrix of the proxies from the previous step. Recently, PLS has been the subject of critical scientific examinations highlighting several shortcomings such as inconsistent parameter estimates for reflective and causal–formative measurement models and a lack of a test for exact overall model fit (Rönkkö and Evermann 2013; Rönkkö et al. 2015, 2016). To overcome these shortcomings, several enhancements have been introduced such as consistent partial least squares (PLSc) to consistently estimate the parameters of reflective and causal–formative measurement models (Dijkstra and Henseler 2015a) and a bootstrap-based test to assess the exact overall model fit (Dijkstra and Henseler 2015b). Moreover, studies provide arguments for and against the use of PLS (e.g., Rigdon 2016) and how PLS can be of value for different types of research (Henseler 2018).

Most recently, in an article in the *Journal of Business Research*, Hair et al. (2020) discuss “CCA” as the measurement model assessment steps in partial least squares structural equation modeling [PLS-SEM, a framework introduced by Hair et al. (2014)]. However, the method they present differs from CCA as originally developed by Schubert et al. (2018) in many regards. First, Hair et al.’s method “can facilitate the assessment of reflective as well as formative measurement models” (Hair et al. 2020, p. 108) and has the purpose “to confirm both reflective and formative measurement models” (p. 101). In contrast, CCA is not meant for assessing measurement models, whether reflective or formative. Instead, its purpose is to assess composite models. Second, Hair et al.’s method “does not require the assessment of fit” (Hair et al. 2020, p. 108), whereas the overall model fit assessment is an essential step in completing a CCA. Third, the methods’ relation to PLS differ strongly. Hair et al.’s method is strongly linked to PLS-SEM; it “is a systematic methodological process for confirming measurement models in PLS-SEM” (Hair et al. 2020, p. 104). In contrast, although CCA *can* use the iterative PLS algorithm as an estimator for the weights and the construct correlation matrix, this is in no way mandatory. There are many estimators for CCA, such as generalized canonical correlation analysis (Kettenring 1971) and generalized structured component analysis (GSCA, Glang 1988; Hwang and Takane 2004). Besides these, in principle also maximum-likelihood, unweighted least squares, and weighted least squares estimators for composite models are conceivable. Consequently, CCA is not tied to PLS. If anything, the iterative PLS algorithm could be regarded as a possible component of CCA. Finally, while there are mathematical proofs and empirical evidence for the efficacy of CCA, such evidence does not exist for Hair et al.’s method. What is worse, one can easily provide evidence against it (see Sect. 5), and the literature has already called the rules of thumb underlying Hair et al.’s method ‘unjustifiable’ (Rönkkö et al. 2016). Overall, CCA and Hair et al.’s method differ in nature, purpose, and efficacy. To prevent researchers confusing CCA with Hair et al.’s method, it is proposed to dub the latter the “method of confirming measurement quality” (MCMQ), a term used by Hair et al. (2020, p. 101). Table 1 summarizes the main differences between CCA and MCMQ.

Against the background of the situation outlined above, the present paper makes a double contribution. First, it describes both CCA and MCMQ and highlights their differences in nature and purpose to prevent researchers from confusing the two. Second, it exposes both methods to a series of scenarios to demonstrate their efficacy in fulfilling their intended purpose. The outcomes show that whereas CCA can indeed distinguish between correctly and incorrectly specified composite models, there are instances in which MCMQ fails to discriminate between correctly and incorrectly specified formative and reflective measurement models, i.e., it tends to confirm measurement models although they are incorrectly specified. These findings imply that (1) if researchers want to assess composite models, they should apply CCA; (2) if researchers want to assess the quality of formative and reflective measurement models, they should use SEM including CFA; and (3) MCMQ could benefit from rethinking, redesigning, and additional investigation.

Table 1 Main differences between confirmatory composite analysis and the method of confirming measurement quality

	Confirmatory composite analysis (CCA, Schuberth et al. 2018)	Method of confirming measurement quality (MCMQ, Hair et al. 2020)
Purpose:	Assessing composite models	Confirming the quality of reflective and formative measurement models
Steps:	Model specification, model identification, model estimation, model assessment	Seven steps to assess reflective measurement models and five steps to assess formative measurement models
Relation to PLS:	Not tied to PLS, but it can serve as an estimator	MCMQ is the evaluation step of PLS-SEM
Role of Fit:	Assessment of model fit is an essential step of CCA	MCMQ does not require the assessment of model fit
Efficacy:	Evidence of its efficacy (mathematical and empirical)	Counterevidence of its efficacy

2 Different types of outer models

To provide solid ground for the discussion, in the following, the three main outer model types in the context of SEM are briefly presented, see e.g., Bollen and Bauldry (2011): (a) the reflective measurement model; (b) the causal–formative measurement model; and (c) the composite model.

The three outer models differ with regard to the relationship of the construct to its observable variables and with regard to the type of construct that represents the abstract concept in the statistical model. Among the three outer models, the following two types of constructs are distinguished: latent variable and emergent variable. While latent variables cannot be inferred with certainty from the data (Borsboom 2008), emergent variables, also referred to as composites, are composed of their observable variables and thus are fully determined by them. It is stressed “that researchers should avoid using the confusing convention of referring to composites as latent variables” (Hardin and Marcoulides 2011, pp. 754–755). In the remainder of this section, it is elaborated on each type of outer model.

Figure 1a displays the reflective measurement model, which is also often referred to as the common factor model. The reflective measurement model assumes that a latent variable (η) causes the observable variables (y) and their interrelations (Jöreskog 1969). Since, the variance in the observable variables can usually not be fully explained by the latent variable, the remaining variance is captured by random errors (ϵ). Typically, these random errors are assumed to be uncorrelated with each other and uncorrelated with the latent variable. As a consequence, the latent variable is the only explanation for the correlations among the observable variables.

The causal–formative measurement model, as illustrated in Fig. 1b, is an alternative type of measurement model (Diamantopoulos 2008). At its core is a set of causal indicators (y), which causally affects the latent variable (η). In doing so, the causal indicators are allowed to be freely correlated. Since the indicators typically

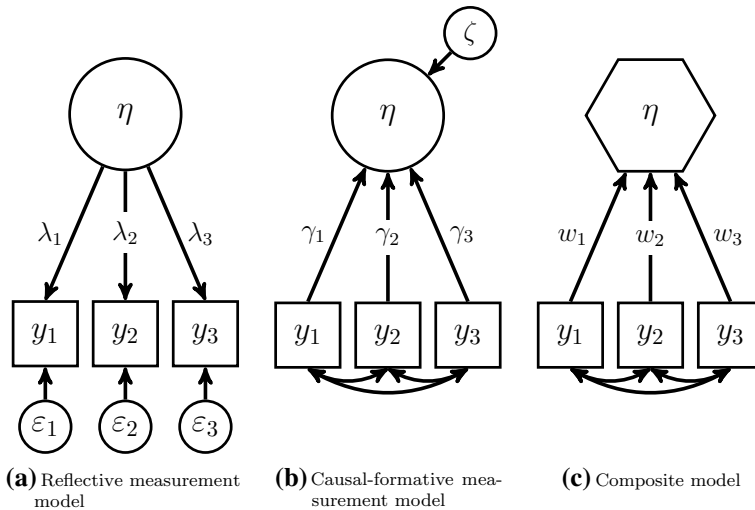


Fig. 1 Three different types of outer models

do not cause all of the variation in the latent variable, an error term (ζ) captures the remaining variance in the construct. It is assumed that the error term is uncorrelated with the causal indicators and that all effects of the causal indicators on other variables in the model are fully mediated by the latent variable.

The composite model is depicted in Fig. 1c. In contrast to the two types of measurement models, the composite model assumes that the construct (η) is fully composed of its observable variables (y), see, e.g., Grace and Bollen (2008). Thus, the construct emerges of its observable variables, wherefore this type of construct is also called emergent variable (Cohen et al. 1990; Cole et al. 1993; Reise 1999; Benitez et al. 2020). As a consequence, there is no error term at the construct level. Similar to the causal–formative measurement model, the observable variables can freely covary.

The choice of the outer model reflects a researcher’s understanding of the world and is not tied to any specific estimator. Every type of outer model requires different rules for model identification and imposes different constraints on the variance–covariance matrix of the observable variables. While the causal–formative measurement model as well as the composite model typically put no constraints on the variance–covariance matrix of the observable variables causing and composing the construct, respectively, the reflective measurement model assumes that the measures are independent when is controlled for the latent variable (Lazarsfeld 1959). However, the former two impose constraints on the covariances between the observable variables connected to a construct and observable indicators connected to other constructs in the model. All these constraints can be exploited in model fit assessment to examine whether the specified model is consistent with the collected data.

3 Confirmatory composite analysis

Only recently, Schuberth et al. (2018) introduced CCA as a subtype of SEM that aims at assessing composite models. Like all forms of SEM, CCA consists of four steps: (1) model specification, (2) model identification, (3) model estimation, and (4) model assessment. The latter is particularly important to assess the estimated model and thus a researcher's theory (Mulaik et al. 1989; Yuan 2005). Approaches that omit one or more of these steps—in particular the last one—would be considered incomplete and hence inapt for fulfilling the purpose of SEM (Rönkkö and Evermann 2013); in other words, “if SEM is used, then model fit testing and assessment is paramount, indeed crucial, and cannot be fudged for the sake of “convenience” or simple intellectual laziness on the part of the investigator” (Barrett 2007, p. 823). The following subsections briefly explain each of the four steps in CCA.

3.1 Specifying composite models

Composite models as typically studied in CCA consist of a set of emergent variables that are allowed to freely covary, although in general, constraints on the variance–covariance matrix of the emergent variables are also conceivable (Dijkstra 2017). Each emergent variable η_j is a linear combination (weighted sum using weights w_{ji}) of I_j observable variables y_{ji} :²

$$\eta_j = \sum_{i=1}^{I_j} w_{ji} \cdot y_{ji} \quad (1)$$

Due to their nature as weighted sums of other variables, emergent variables are essentially prescriptions for dimension reduction (Dijkstra and Henseler 2011). The analogy between CCA and CFA is obvious and intended: Whereas a CFA usually studies a set of interrelated latent variables, a CCA examines a set of interrelated emergent variables.

Figure 2 shows an exemplary composite model consisting of three interrelated emergent variables, each of which is composed of three observable variables. This composite model allows the observable variables of each emergent variable to be freely correlated as highlighted by the double-headed arrows among the observable variables belonging to one emergent variable. Similarly, all emergent variables are allowed to be freely correlated. Finally, the emergent variables fully convey the correlations between the observable variables of two different emergent variables.

² Although the study at hand focus only emergent variables composed by observable variables, in general, emergent variables can be also made up by latent or other emergent variables, see for example van Riel et al. (2017) and Schuberth et al. (2020).

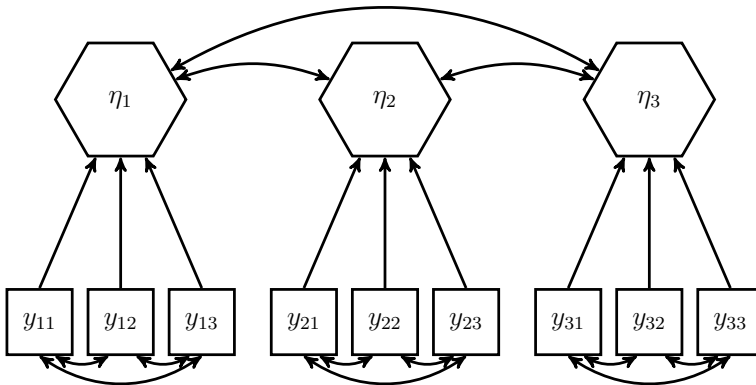


Fig. 2 A composite model consisting of three correlated emergent variables

3.2 Identifying composite models

Model identification plays a preponderant role in SEM including CFA and CCA. Once the model is specified, it needs to be ensured that the parameters of the specified model can be uniquely retrieved from the variance–covariance matrix of the observable variables, i.e., it needs to be assessed whether the model is identified. Interpreting parameter estimates of underidentified models leads to questionable conclusions, as several sets of parameters exist that satisfy the constraints of the model-implied variance–covariance matrix, i.e., several parameter sets lead to the same model-implied variance–covariance matrix.

Similar to CFA, a necessary condition for the identification of a composite model is to fix the scale of each emergent variable in the model. This can be done by either fixing one weight per emergent variable or fixing the variance of each emergent variable. Typically, the variance of each emergent variable is fixed to one by employing properly scaled weights. If this approach is applied, the sign of each weight vector of every block also needs to be determined, similar to a reflective measurement model if the variance of the latent variable is fixed to one. Moreover, no emergent variable is allowed to be isolated, i.e., each emergent variable must be correlated with at least one emergent variable in the model.³ Otherwise, the model-implied covariances between the observable variables of an isolated emergent variable and the observable variables of the remaining emergent variables are all equal to zero, which implies for the isolated emergent variable an infinite number of weight sets that satisfies the scaling condition.

³ While the original article on CCA also allows for observable variables that are correlated with the emergent variable, for simplicity, the study at hand only focuses on models containing correlated emergent variables.

3.3 Estimating composite models

As common in SEM, CCA deliberately distinguishes between the model and the estimator. Although it is suggestive to employ a composite-based estimator such as the iterative PLS algorithms to estimate the weights and the emergent variables' correlations of the composite models (Henseler 2016), in general, other estimators such as GSCA or approaches to generalized canonical correlation analysis can be employed. The guiding criteria in the decision in favor of a certain estimator are its statistical properties and its implied assumptions about the underlying model and population, i.e., the estimator must conform the assumptions imposed by the composite model. Predominantly, researchers should favor unbiased and/or consistent estimators over biased/inconsistent estimators. A consistent estimator produces estimates that converge in probability towards their population counterpart; an unbiased estimator produces estimates whose expected value equals the population counterpart. Purely practical aspects such as computation time and convergence behavior tend to play a minor role in the selection of estimators.

3.4 Assessing composite models

Similar to CFA and SEM, model assessment is a crucial step in CCA. In doing so, the composite model is usually assessed globally and locally. While the global model assessment examines the model as a whole, local model assessment investigates each emergent variable separately including the relationship to its observable variables.

The global model assessment, also known as overall model fit assessment, is a crucial step in SEM as well as CCA and its importance is acknowledged across various disciplines and emphasized in literally every textbook on SEM (e.g., Schumacker and Lomax 2016; Kline 2015). The most well-known ways to assess a model's fit are: (i) fit indices and (ii) the test for exact overall model fit. The former quantifies the misfit on a continuum and the decision about whether the model acceptably fits the collected data is usually based on heuristic rules. In contrast, the latter assesses the exact fit of the overall model by means of statistical testing. In doing so, model fit assessment and decision making are based on a p -value.

In general, the test for exact overall model fit compares the sample variance–covariance matrix of the observable variables to their model-implied counterpart. In doing so, the test assesses the null hypothesis that the model-implied variance–covariance matrix based on the population parameters equals the population variance–covariance matrix of the observable variables: $H_0 : \Sigma(\theta) = \Sigma$. In other words, the test for exact overall model fit examines whether it is plausible that the world functions as described by the specified model.

Various tests are available for this endeavor depending on the statistical properties of the employed estimator and the assumptions about the population. To keep the number of needed assumptions low, Schuberth et al. (2018) proposed employing a nonparametric test for exact overall model fit assessment that obtains the distribution of the discrepancy measure under the null hypothesis through the bootstrap

(Beran and Srivastava 1985). This test has the advantage that as long as the selected estimator produces consistent estimates for the population parameters, the asymptotic distribution of the test statistic is not tied to a specific estimator.⁴ The asymptotic properties of the test are mathematically derived, and its finite sample performance for composite models has been demonstrated (Schubert et al. 2018).⁵

Once the model shows an acceptable model fit, the researcher can proceed with the local assessment of the composite model, i.e., each emergent variable with its respective observable variables is considered separately. In doing so, the significance of the estimated weights, their sign, and size are examined and matched with a researcher's theory. Similarly, the correlations among the emergent variables are matched with the expectations of a researcher. Moreover, if the weight estimates are subject to multicollinearity, it is recommended to inspect the correlational patterns of the observable variables forming an emergent variable.

4 Hair's method of confirming measurement quality

Hair et al. (2020) introduce MCMQ as the measurement model assessment step in PLS-SEM (Hair et al. 2014). MCMQ entails different evaluation steps, depending on whether analysts define their measurement models as reflective or formative. The next two subsections briefly describe the evaluation steps for reflective and formative measurement models and add some critical reflections where needed. For a more elaborate treatment, we refer to Hair et al. (2020).

4.1 Assessing reflective measurement models using MCMQ

Once the PLS algorithm has been executed, MCMQ requires seven steps to evaluate a reflective measurement model (see Table 2 of Hair et al. 2020):

1. *Estimate of loadings and significance*

In PLS-SEM, *loading estimates* are correlations between the construct scores and the corresponding observable variables. According to Hair et al. (2020), the loading estimates must be significantly different from zero and have a value of

⁴ In addition, some trivial assumptions such as independent and identically distributed (i.i.d.) observable variables must hold; see Beran and Srivastava (1985) for more details.

⁵ Against this background, Hair et al.'s pejorative remark that a "simulation study *attempted* to illustrate the performance of bootstrap-based tests and discrepancy measures" (2020, p. 103, emphasis added by the author) is unfounded. It is also not understandable why a nonparametric test should be discarded only because a simulation has shown that it performs well under normality. Finally, the fact that the iterative PLS algorithm is employed in the case of small sample sizes does not reduce the importance of model fit assessment in CCA; rather, it calls for more research on model fit assessment in such situations; otherwise, researchers will face statistical tests with a low statistical power which leads questionable conclusions. Additionally, it is well-known in the PLS literature that small sample sizes are problematic, see e.g., Rigdon (2016, p. 600) who states that "Yes, PLS path modeling will produce parameter estimates even when sample size is very small, but reviewers and editors can be expected to question the value of those estimates, beyond simple data description".

0.708 or greater. However, Hair et al. (2020) fail to mention that, if interpreted as loadings of a reflective measurement model, the loading estimates in PLS-SEM tend to be strongly upward biased (McDonald 1996).

2. *Indicator reliability (items)*

The concept of indicator reliability is strongly connected to the reflective measurement model and indicates how much of the variance in an indicator is explained by its underlying latent variable. In PLS-SEM, the *indicator reliability* is calculated as the squared loading estimate. Hair et al. (2020) do not provide any further explanations or any threshold level for the indicator reliability, so it remains unclear how and for what purpose researchers should employ it. Since the loading estimates in PLS-SEM are upwardly biased for reflective measurement models, so are the indicator reliabilities.

3. *Composite reliability (construct)*

According to Hair et al. (2020), the *composite reliability* of construct scores should range from 0.7 to 0.95. The coefficient of composite reliability is inherently tied to the reflective measurement model and assumes that the measures of a latent variable are prone to random measurement errors. It equals the ratio of the variance of the latent variable and the variance of a linear combination of the latent variable's indicators (Jöreskog 1971). Since in PLS-SEM, the composite reliability is calculated based on upwardly biased loadings, composite reliability will be inflated. For instance, a construct measured by two observable variables with an indicator reliability of 0.28 each will yield a composite reliability greater than 0.7, although the true reliability of the construct scores is 0.4375. Researchers relying on MCMQ may thus mistakenly believe that their measures are reliable, whereas in fact, the reliability of their measures is far below the threshold of 0.7 that is recommended by Nunnally and Bernstein (1994). A more detailed presentation including negative consequences is given in the "Appendix".

4. *Average variance extracted (AVE)*

According to Fornell and Larcker (1981), the *average variance extracted (AVE)* should exceed 0.5. The AVE assumes a reflective measurement model and indicates how much of the variance in the indicators is explained by the underlying latent variable (Fornell and Larcker 1981). In PLS-SEM, the AVE is determined by averaging the indicator reliabilities (Hair et al. 2020). Since the indicator reliabilities are upwardly biased, so is the AVE. For instance, a construct that is measured by two observable variables with an indicator reliability of 0.28 each will yield an AVE clearly higher than 0.5 in PLS-SEM, although the true AVE would only be 0.28. Researchers relying on MCMQ may thus mistakenly conclude that their measures possess convergent validity, whereas in fact, they do not.

5. *Discriminant validity—HTMT*

Whether two constructs can be statistically discriminated should be assessed by means of the heterotrait-monotrait ratio of correlations (HTMT, Henseler et al. 2015). The HTMT assumes a reflective measurement model and provides a consistent estimate for the correlation between two latent variables if the reflective measurement model is tau-equivalent, see the Appendix of Henseler et al. (2015). Since the HTMT is only based on the correlation matrix of the observable

variables, its values remain untouched by PLS-SEM. However, it is noted that the HTMT assumes an underlying reflective measurement model for which PLS-SEM produces inconsistent estimates, regardless of whether Mode A or B is employed to calculate the weights.

6. *Nomological validity*

The *nomological validity* of constructs can be assessed by means of verifying that the correlational pattern of the constructs is in line with the expected relationships based on extant theory (Hair et al. 2020). However, in PLS-SEM, for reflective measurement models, the correlations among latent variables are biased, and thus, conclusions about the nomological validity are questionable.

7. *Predictive validity*

Finally, *predictive validity* “assesses the extent to which a construct score predicts scores on some criterion measure” (Hair et al. 2020, p. 105). If a researcher’s goal is pure prediction, he is of course not tied rules of confirmatory research.

Remarkably, the assessment of overall model fit is not part of the evaluation steps. As a consequence, MCMQ does not assess whether the specified model is consistent with the collected data. Moreover, several of the assessment steps involve assessment criteria that assume a reflective measurement model for which in PLS-SEM biased estimates are produced. Therefore, conclusions based on these criteria are questionable.

4.2 Assessing formative measurement models using MCMQ

For formative measurement models, MCMQ entails five evaluation steps (see Table 3 of Hair et al. 2020):

1. *Convergent validity—redundancy*

The latent variable’s coefficient of determination (R^2) refers to *convergent validity*. It quantifies the extent of the latent variable’s variance that is explained by its formative indicators. Instead of the R^2 , one can also assess its square root, which represents the effect of an optimal linear combination of formative indicators on the latent variable. As the former and the latter are typically standardized in PLS-SEM, this effect equals the path coefficient between the two. Hair et al. (2020) postulate a path coefficient of at least 0.7, which corresponds to a minimal R^2 value of 0.49. However, this rule of thumb neglects that in PLS-SEM the path coefficients are biased if the dependent or the independent variable (or both) is latent. For instance, if the latent variable has two reflective indicators with a reliability of 0.28 each (which will remain unnoticed by MCMQ; see the explanation above), a path coefficient of 0.7 is impossible.

2. *Indicator multicollinearity*

According to Hair et al. (2020), researchers must assess *indicator multicollinearity* because it “creates problems with formative measurement models” (p. 105). Variance inflation factors (VIFs) should not exceed 3, and bivariate correlations should be lower than 0.5 (Hair et al. 2020). Notably, bivariate correlations can

exceed 0.8 without necessarily yielding a VIF above 3. At the same time, VIF values greater than 10 can occur even if no bivariate correlation exceeds 0.5. Consequently, this assessment step is likely to often render measurement models ‘problematic’. Moreover, multicollinearity is a characteristic of the sample and not a problem of the underlying model.

3. *Size and significance of indicator weights*

Convergent validity and absence of multicollinearity are prerequisites for the next step, the assessment of the *size and significance of indicator weights* (Hair et al. 2020). The weights of the formative indicators should be significant; indicators with insignificant weights may or may not be discarded (Hair et al. 2020).

4. *Contribution of indicators (size and significance of loadings)*

The loadings of formative indicators should be significantly different from zero and exceed 0.5; otherwise, the researcher can discard or retain formative indicators (Hair et al. 2020).

5. *Predictive validity*

Finally, the assessment of formatively measured constructs’ *predictive validity* is analogous to that of reflectively measured constructs (Hair et al. 2020). Therefore, it is only repeated that researchers who are interested in pure predictive research do not have to follow the rules for rigorous confirmatory research.

Again, the assessment of overall model fit is not part of the evaluation steps. As a consequence, MCMQ does not inform researchers on whether a specified formative measurement model is consistent with the collected data.

5 Assessment of confirmatory composite analysis and Hair’s method of confirming measurement quality by means of a scenario analysis

The inventors of CCA and MCMQ make different claims about what their methods are suitable for: CCA is meant to assess composite models (Schuberth et al. 2018), and MCMQ’s objective is the “confirmation of the measurement models” (Hair et al. 2020, p. 103), whether reflective or formative. However, whereas Schuberth et al. (2018) demonstrated the efficacy of CCA by means of a Monte Carlo simulation, MCMQ has not been exposed to any rigorous examination so far. Although Hair et al. (2020, p. 108) state that MCMQ “is a superior approach”, they do not provide any evidence. Consequently, two fundamental questions about the absolute and relative efficacy of MCMQ arise: Does MCMQ actually detect problematic models? And if so, is it superior to SEM including CCA and CFA? A scenario analysis can help to answer these questions.

5.1 Setup of the scenario analysis

Scenario analysis allows us to illustrate how well different methods retrieve parameters from a given population; “estimation” is thus not done on the basis of a sample.

Hence, parameters are rather retrieved than estimated. This avoids distributional assumptions and uncertainty introduced by sampling.

To examine to what extent SEM, including CCA and CFA, and MCMQ are able to assess the quality of a model, the methods are exposed to six scenarios. Each scenario is a combination of a population model and a specified model of a fictitious researcher. In the Scenarios 1a, 2a, and 3a, the researcher's model is correctly specified and the methods employed to assess the model are expected not to indicate any problems. In contrast, in the three Scenarios 1b, 2b, and 3b, the studied population functions differently than the way the researcher thinks it does. In these situations, the researcher's model is incorrectly specified. Hence, the methods employed to assess the model are expected to sound the alarm and indicate that the model is problematic.

In Scenarios 1a and 1b, a researcher thinks that the population functions according to a multiple-indicators, multiple-causes (MIMIC, Jöreskog and Goldberger 1975) model. A MIMIC model entails that a latent variable is measured with both reflective and causal-formative indicators, i.e., it combines causal-formative and reflective measurement. In our case, there are six indicators, three of which are considered reflective and the other three to be causal-formative. Whereas in Scenario 1a, the population indeed functions according to a MIMIC model, in Scenario 1b, the population functions according to a different model. Since the purpose of MCMQ is to confirm formative and reflective measurement models, a MIMIC model is a formidable test case for MCMQ. Both MCMQ and classical SEM are used to assess the MIMIC model. Obviously, in Scenario 1a, the methods should not reject the researcher's model, whereas in Scenario 1b, the methods should make the researcher aware that his or her model is problematic.

In Scenarios 2a and 2b, a researcher thinks that the population functions according to a composite model. Concretely, the specified model consists of two correlated emergent variables, each composed of three observable variables. Whereas in Scenario 2a, the population indeed functions according to the composite model, in Scenario 2b, the population functions according to a different model.

Since CCA serves to assess composite models, these scenarios allow us to illustrate its performance. Notably, Hair et al. regard MCMQ "as a separate approach to confirming linear composite constructs in measurement models" (p. 104). For them, "reflective measurement models are composite latent constructs" (Hair et al. 2020, p. 104), and "[f]ormative composite measurement models are linear combinations of a set of indicators that form the construct" (p. 105). Moreover, MCMQ "enables researchers to develop and validate measures within a nomological network. Each composite, therefore, must relate to at least one other composite. Hence, the validity of a composite depends on the nomological network in which it is embedded." (Hair et al. 2020, p. 103) These statements suggest that MCMQ might be suitable for assessing composite models as analyzed in CCA. Therefore, MCMQ is also employed in Scenarios 2a and 2b. In Scenario 2a, both CCA and MCMQ should not sound alarm, whereas in Scenario 2b, the methods should alert the researcher that the specified model is problematic.

Against the recent history of PLS-SEM, and especially since Rigdon (2012), it is suggested that in PLS-SEM reflective and formative measurement refer to a model

where the emergent variable is built by correlation weights (Mode A) and regression weights (Mode B), respectively. As MCMQ is applied to PLS-SEM, a further composite model is investigated in the last two scenarios. In contrast to Scenario 2a, in Scenario 3a, a population is considered where the second emergent variable is built by correlation weights. To generate the population weights for this emergent variable, an approach suggested by Cho and Choi (2020) is applied. Similar to Scenarios 2a and 2b, in Scenario 3a, the specified model matches the underlying population, while in Scenario 3b, the model is misspecified. It is noted that the model where an emergent variable is built by correlation weights is nested within the composite model analyzed in CCA. Hence, also CCA is employed in Scenarios 3a and 3b.

Similar to Scenarios 1 and 2, in Scenario 3a, the specified model is correct, while in Scenario 3b, the population does not function according to the specified model. Consequently, in Scenario 3a, CCA and MCMQ should indicate no problems in the specified model. In contrast, in Scenario 3b, the specified model is incorrect and this should also be reflected by CCA and MCMQ.

Table 2 provides an overview of the six scenarios. It includes the population models, their correlation matrices, and the models as specified by the fictitious researcher. In the Scenarios 3a and 3b, the second construct η_2 is deliberately displayed by a hexagon to highlight the fact that it is an emergent variable, although the arrows point from the construct to the indicators. This implies that the observable variables of the second emergent variable are also allowed to be freely correlated, and it also emphasizes the difference of that special type of emergent variable with the classical reflective measurement model known from SEM. All values are rounded to the second decimal place.

MCMQ was exposed to all six scenarios. Since MCMQ is the evaluation step in PLS-SEM, PLS-SEM was conducted as outlined in a current primer on that framework (concretely, Hair et al. 2017). Subsequently, the evaluation steps proposed by Hair et al. (2020) were performed and the corresponding conclusions were drawn. Three evaluation steps were omitted. First, discriminant validity was not assessed because it is not applicable to models with less than two reflectively measured constructs. Second, nomological validity was not assessed in Scenarios 1a and 1b, as only one construct, i.e., the latent variable, is available. Third, although PLS can be employed for predictive modeling (Cepeda Carrión et al. 2016), predictive validity was not considered because it is not a mandatory part of confirmatory research. It is well-known that in causal research, i.e., explanatory and confirmatory research, the correctness of the specified model is substantial, while “‘wrong’ model can sometimes predict better than the correct one” (Shmueli 2010, p. 6). Moreover, similar guidelines (such as, for instance, in Hair et al. 2017) do not include predictive validity.⁶

SEM and CCA were performed along their four constituting steps: model specification, identification, estimation, and assessment. To keep the difference between the different methods as small as possible, PLS was employed as an estimator. To

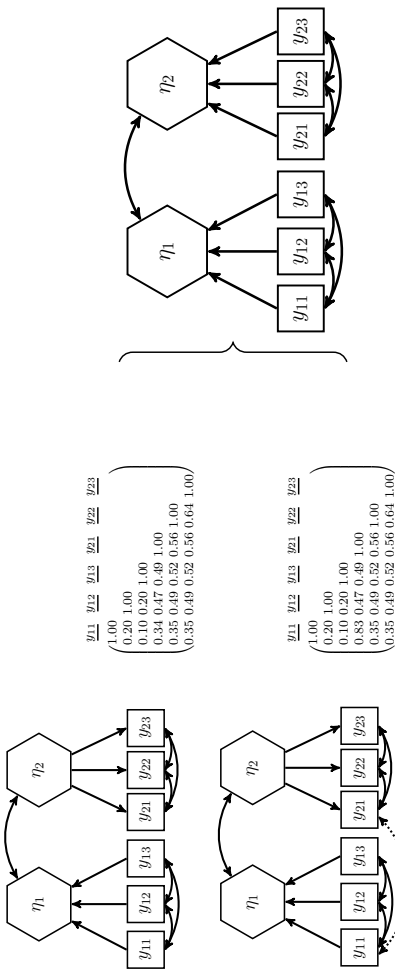
⁶ It would have been easy to add a moderately highly correlated dependent variable to each scenario to yield predictive validity. The study at hand refrained from this exercise to keep the models simple.

Table 2 The six scenarios

Scenario	Description	Correlation matrix	Population model	Specified model
Scenario 1a	Population functions according to a MIMIC model, and the analyst studies a model with formative and reflective measurement models. The analyst's model is correct.	$\begin{pmatrix} \eta_{11} & \eta_{12} & \eta_{13} & \eta_{21} & \eta_{22} & \eta_{23} \\ 1.00 & & & & & \\ 0.20 & 1.00 & & & & \\ 0.29 & 0.40 & 0.42 & 1.00 & & \\ 0.33 & 0.46 & 0.48 & 0.56 & 1.00 & \\ 0.33 & 0.46 & 0.48 & 0.56 & 0.64 & 1.00 \end{pmatrix}$		
Scenario 1b	Population functions according to a model different from a MIMIC model, yet the analyst studies a model with formative and reflective measurement models. The analyst's model is incorrect.	$\begin{pmatrix} \eta_{11} & \eta_{12} & \eta_{13} & \eta_{21} & \eta_{22} & \eta_{23} \\ 1.00 & & & & & \\ 0.20 & 1.00 & & & & \\ 0.40 & 0.20 & 0.42 & 1.00 & & \\ 0.33 & 0.46 & 0.48 & 0.56 & 1.00 & \\ 0.33 & 0.46 & 0.48 & 0.56 & 0.64 & 1.00 \end{pmatrix}$		
Scenario 2a	Population functions according to a composite model, and the analyst studies a model with interrelated emergent variables. The analyst's model is correct.	$\begin{pmatrix} \eta_{11} & \eta_{12} & \eta_{13} & \eta_{21} & \eta_{22} & \eta_{23} \\ 1.00 & & & & & \\ 0.15 & 1.00 & & & & \\ 0.30 & 0.30 & 1.00 & & & \\ 0.15 & 0.25 & 0.22 & 1.00 & & \\ 0.22 & 0.37 & 0.32 & 0.20 & 1.00 & \\ 0.24 & 0.39 & 0.34 & 0.00 & 0.40 & 1.00 \end{pmatrix}$		
Scenario 2b	Population functions according to a model different from a composite model, yet the analyst studies a model with interrelated emergent variables. The analyst's model is incorrect.	$\begin{pmatrix} \eta_{11} & \eta_{12} & \eta_{13} & \eta_{21} & \eta_{22} & \eta_{23} \\ 1.00 & & & & & \\ 0.15 & 1.00 & & & & \\ 0.30 & 0.30 & 1.00 & & & \\ 0.64 & 0.25 & 0.22 & 1.00 & & \\ 0.22 & 0.37 & 0.32 & 0.20 & 1.00 & \\ 0.24 & 0.39 & 0.34 & 0.00 & 0.40 & 1.00 \end{pmatrix}$		

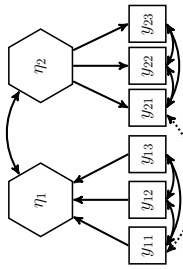
Table 2 (continued)

Scenario 3a Population functions according to a composite model, and the analyst studies a composite model. **The analyst's model is correct.**



$$\begin{pmatrix} 1.00 & 0.20 & 0.10 & 0.35 & 0.49 & 0.52 & 0.56 & 0.64 & 1.00 \\ 0.20 & 1.00 & 0.34 & 0.47 & 0.49 & 1.00 & 0.35 & 0.49 & 0.52 & 0.56 & 1.00 \\ 0.10 & 0.34 & 1.00 & 0.35 & 0.49 & 0.52 & 0.56 & 0.64 & 1.00 \\ 0.35 & 0.49 & 0.35 & 1.00 & 0.20 & 0.10 & 0.83 & 0.47 & 0.49 & 1.00 \\ 0.49 & 0.47 & 0.49 & 0.83 & 1.00 & 0.10 & 0.20 & 1.00 & 0.35 & 0.49 & 0.52 & 0.56 & 1.00 \\ 0.52 & 0.56 & 0.52 & 0.56 & 0.64 & 1.00 & 0.35 & 0.49 & 0.52 & 0.56 & 1.00 \end{pmatrix}$$

Scenario 3b Population functions according to a model different from the composite model, yet the analyst studies a composite model. **The analyst's model is incorrect.**



$$\begin{pmatrix} 1.00 & 0.20 & 0.10 & 0.35 & 0.49 & 0.52 & 0.56 & 0.64 & 1.00 \\ 0.20 & 1.00 & 0.34 & 0.47 & 0.49 & 1.00 & 0.35 & 0.49 & 0.52 & 0.56 & 1.00 \\ 0.10 & 0.34 & 1.00 & 0.35 & 0.49 & 0.52 & 0.56 & 0.64 & 1.00 \\ 0.35 & 0.49 & 0.35 & 1.00 & 0.20 & 0.10 & 0.83 & 0.47 & 0.49 & 1.00 \\ 0.49 & 0.47 & 0.49 & 0.83 & 1.00 & 0.10 & 0.20 & 1.00 & 0.35 & 0.49 & 0.52 & 0.56 & 1.00 \\ 0.52 & 0.56 & 0.52 & 0.56 & 0.64 & 1.00 & 0.35 & 0.49 & 0.52 & 0.56 & 1.00 \end{pmatrix}$$

correct the parameters of the reflective measurement model for attenuation in Scenario 1a and 1b, PLSc was applied for SEM. To obtain the parameters in Scenarios 2a and 2b, the iterative PLS algorithm was employed using mode B for both emergent variables. Similar in Scenarios 3a and 3b, Mode B was used for the first emergent variable and Mode A for the second one.⁷ Finally, three coefficients helped to assess the overall fit of the models: (1) the standardized root mean square residual (SRMR, Hu and Bentler 1998) and (2) the geodesic discrepancy (d_G) as measures of misfit (Dijkstra and Henseler 2015a) and (3) the normed fit index (NFI, Bentler and Bonett 1980) as a relative fit measure. For the latter, the model-implied correlation matrix of the competing model equals a unit matrix.

5.2 Results and conclusion

Table 3 presents the parameters of all models involved: both the parameters of the population models and the retrieved parameters⁸ of the specified models. In Scenario 1a, as expected and well-known in the literature (e.g., Dijkstra 1981), the iterative PLS algorithm produces Fisher-consistent estimates for the composite model but not for the reflective measurement model. To retrieve the population parameters correctly from reflective measurement models, PLSc is used as an estimator in SEM. In contrast, PLS-SEM and thus MCMQ rely on traditional PLS without a correction for attenuation, and as a consequence it does not retrieve the population parameters. This is a deliberate choice of the inventors of PLS-SEM because in their view, PLSc “adds very little to the body of knowledge” (Hair et al. 2019b, p. 570). In Scenario 2a, which purely deals with emergent variables, the iterative PLS algorithm delivers the true parameter values as a result. Similarly, in Scenario 3a, which contains the *special* emergent variable formed by correlation weights, the PLS algorithm using Mode B for the first emergent variable and Mode A for the second emergent variable retrieves the population parameters. Consequently, both CCA and MCMQ are based on the population parameters in Scenarios 2a and 3a. In Scenarios 1b, 2b, and 3b, where the specified model does not match the population model, all estimators provide parameter values that deviate from the population values. As a consequence, CCA, SEM, and MCMQ are all based on incorrect parameters, which shows that a misspecification of the model generally prevents the fictitious researcher from obtaining an accurate understanding of the world.

For Scenarios 1a and 1b, the results of MCMQ and SEM are displayed in Table 4. MCMQ confirms the quality of the specified model in both scenarios, and none of the proposed steps could detect the problem of the specification in Scenario 1b. To assess the outer model of the first construct η_1 , MCMQ’s rules for formative measurement models are followed, which lead, for both scenarios, to the same results.

⁷ It is noted that emergent variables made up by correlation weights are consistently estimated by both Mode A and Mode B using the iterative PLS algorithm, as this type of composite model is a special case of the model where the emergent variable is built by regression weights (Cho and Choi 2020).

⁸ To retrieve the parameters and calculate the statistics reported in the following, the R package cSEM is used (Rademaker and Schubert 2020).

Table 3 Population and retrieved parameters for the six scenarios

Parameter	Scenario 1a			Scenario 1b			Scenario 2a			Scenario 2b			Scenario 3a			Scenario 3b		
	True	MCMQ	SEM	True	MCMQ	SEM	True	MCMQ	SEM	True	MCMQ	CCA	True	MCMQ	CCA	True	MCMQ	CCA
w_{11}	0.347	0.347	0.347	0.347	0.606	0.606	0.274	0.274	0.274	0.274	0.274	0.770	0.347	0.347	0.347	0.347	0.347	0.576
w_{12}	0.520	0.520	0.520	0.520	0.387	0.387	0.658	0.658	0.658	0.658	0.658	0.426	0.520	0.520	0.520	0.520	0.520	0.406
w_{13}	0.607	0.607	0.607	0.607	0.500	0.500	0.439	0.439	0.439	0.439	0.439	0.148	0.607	0.607	0.607	0.607	0.607	0.516
w_{21}	-	0.357	0.367	-	0.480	0.480	0.422	0.422	0.422	0.422	0.422	0.808	0.379	0.379	0.379	0.379	0.379	0.479
w_{22}	-	0.408	0.408	-	0.347	0.347	0.411	0.411	0.411	0.411	0.411	0.153	0.398	0.398	0.398	0.398	0.398	0.348
w_{23}	-	0.408	0.408	-	0.347	0.347	0.347	0.617	0.617	0.617	0.617	0.466	0.398	0.398	0.398	0.398	0.398	0.348
β/ρ	0.800	0.722	0.800	0.800	0.834	0.911	0.600	0.600	0.600	0.600	0.600	0.753	0.800	0.800	0.800	0.800	0.800	0.908
λ_{11}	-	0.512	0.512	-	0.734	0.734	-	0.504	0.504	0.504	0.504	0.878	-	0.512	0.512	-	0.709	
λ_{12}	-	0.711	0.711	-	0.609	0.609	-	0.831	0.831	0.831	0.831	0.586	-	0.711	0.711	-	0.624	
λ_{13}	-	0.746	0.746	-	0.638	0.638	-	0.718	0.718	0.718	0.718	0.506	-	0.746	0.746	-	0.655	
λ_{21}	0.700	0.814	0.700	0.700	0.869	0.931	-	0.504	0.504	0.504	0.838	0.838	-	0.825	0.825	-	0.869	
λ_{22}	0.800	0.869	0.800	0.800	0.839	0.674	-	0.742	0.742	0.742	0.501	0.501	-	0.864	0.864	-	0.839	
λ_{23}	0.800	0.869	0.800	0.800	0.839	0.674	-	0.782	0.782	0.782	0.528	0.528	-	0.864	0.864	-	0.839	

Table 4 Results of MCMQ and SEM for Scenarios 1a and 1b

Scenario	The analyst considers a model with causal–formative and reflective measurement	Scenario 1a:	Scenario 1b:
	The population functions according to a MIMIC model	The population functions according to a model different from the MIMIC model	
Situation	The analyst's model is correct	The analyst's model is incorrect	
MCMQ results	<p>Rules for formative measurement models:</p> <ul style="list-style-type: none"> ✓ Convergent validity ($\beta = 0.722 > 0.70$) ✓ Indicator multicollinearity (largest VIF: 1.369 < 3; largest correlation: 0.200 < 0.5) ✓ Size and significance of weights (smallest weight: 0.347) ✓ Size and significance of loadings (smallest loading: 0.512 > 0.5) ○ Predictive validity <p>Rules for reflective measurement models:</p> <ul style="list-style-type: none"> ✓ Size and significance of loadings (smallest loading: 0.814 > 0.708) ✓ Indicator reliability (smallest indicator reliability: 0.663) ✓ Composite reliability (Cronbach's $\alpha = 0.810 > 0.7$; CR = 0.887 > 0.7) ✓ Average variance extracted (AVE): 0.725 > 0.5) ○ Discriminant validity ○ Nomological validity ○ Predictive validity 	<p>Rules for formative measurement models:</p> <ul style="list-style-type: none"> ✓ Convergent validity ($\beta = 0.834 > 0.70$) ✓ Indicator multicollinearity (largest VIF: 1.369 < 3; largest correlation: 0.200 < 0.5) ✓ Size and significance of weights (smallest weight: 0.387) ✓ Size and significance of loadings (smallest loading: 0.609 > 0.5) ○ Predictive validity <p>Rules for reflective measurement models:</p> <ul style="list-style-type: none"> ✓ Size and significance of loadings (smallest loading: 0.839 > 0.708) ✓ Indicator reliability (smallest indicator reliability: 0.703) ✓ Composite reliability (Cronbach's $\alpha = 0.810 > 0.7$; CR = 0.886 > 0.7) ✓ Average variance extracted (AVE): 0.721 > 0.5) ○ Discriminant validity ○ Nomological validity ○ Predictive validity 	

Table 4 (continued)

Scenario	The analyst considers a model with causal–formative and reflective measurement	
	Scenario 1a:	Scenario 1b:
	The population functions according to a MIMIC model	The population functions according to a model different from the MIMIC model
Situation	The analyst's model is correct	The analyst's model is incorrect
SEM results	Model fit criteria: ✓ SRMR = 0.000 ✓ $d_G = 0.000$ ✓ NFI = 1.000	Model fit criteria: ✗ SRMR = 0.088 ✗ $d_G = 0.123$ ✗ NFI = 0.784
Conclusion	MCMQ confirms measurement quality, and CFA/SEM does not reject the model	MCMQ confirms measurement quality, but CFA/SEM rejects the model

✓: the criterion indicates no problem; ✗: the criterion indicates a problem; O: the criterion is not applicable in this scenario

As suggested by Hair et al. (2020), convergent validity is established since the path coefficient between the emergent variable and the latent variable is above 0.7 (see Table 3). Moreover, multicollinearity is not regarded as an issue since the correlations among the three causal indicators are below 0.5 (see the correlation matrix of Scenarios 1a and 1b in Table 2). Similarly, the VIFs are all below the suggested threshold of 3.⁹ The retrieved weights for the causal–formative indicators are all sizable (see Table 3) and range from 0.347 to 0.607.¹⁰ Similarly, the loadings of the causal indicators range from 0.512 to 0.746 (see Table 3) and thus are all above the recommended threshold of 0.5.¹⁰ To assess the reflective measurement model of the latent variable, MCMQ’s quality criteria for reflective measurement models are applied. The criteria do not indicate any problems for either Scenario 1a or 1b. The corresponding loadings are all above 0.7 (see Table 3).¹⁰ Since there is no threshold mentioned by Hair et al. (2020), indicator reliabilities ranging from 0.663 to 0.755 are regarded as sufficient. For both scenarios, the composite reliability measures are also between the proposed thresholds of 0.70 and 0.95: Scenario 1a: Cronbach’s $\alpha = 0.810$ and Jöreskog’s $\rho = 0.887$, and Scenario 1b: Cronbach’s $\alpha = 0.810$ and Jöreskog’s $\rho = 0.886$. Similarly, the AVE is 0.725 in Scenario 1a and 0.721 in Scenario 1b and thus above the proposed threshold of 0.5. As a consequence, MCMQ does not alert the researcher that his or her model of Scenario 1b is misspecified.

Considering the model fit assessment criteria in Table 4, they correctly indicate no problems in the specified model of Scenario 1a. The SRMR as well as the geodesic distance are both zero, while the NFI equals 1. In contrast, for Scenario 1b, both the SRMR and the geodesic distance substantially exceed 0. Moreover, the normed fit index (NFI) is clearly below 1. As a consequence, the researcher is alerted that his or her model is misspecified.

Similar to the assessment of the reflective and causal–formative measurement models, Table 5 shows that MCMQ’s quality criteria for formative measurement models do not indicate any problems with the specified composite models for both Scenarios 2a and 2b. Since there are no reflective measures of the two emergent variables, convergent validity has not been assessed. Considering the correlations among the observable variables (see the correlation matrix of Scenarios 2a and 2b in Table 2) and the VIF values for the weights of the two emergent variables, multicollinearity is not an issue.¹¹ Moreover, all weights are sizable, and the loadings are above the proposed threshold of 0.5 (see Table 3).¹⁰ As a consequence, MCMQ does not alert the researcher that his or her composite model in Scenario 2b is misspecified.

In contrast, the proposed model fit criteria in CCA do not reject the specified composite model of Scenario 2a and correctly indicate problems in the composite

⁹ In both scenarios, the VIF values for the weights of the causal indicators are as follows: w_{11} : 1.264; w_{12} : 1.369; and w_{13} : 1.264.

¹⁰ Besides considering the size of the parameters, Hair et al. (2020) propose to examine the significance of parameter estimates. In this scenario analysis the parameters are directly retrieved from the population and thus the parameters are not estimated based on a sample. Hence, there is no need to assess the significance of a parameter estimate in order to draw conclusions about its counterpart in the population.

¹¹ In both scenarios, the VIF values of the weights are as follows: w_{11} : 1.442, w_{12} : 1.442, w_{13} : 1.655, w_{21} : 1.279, w_{22} : 1.809, and w_{23} : 1.690

Table 5 Results of MCMQ and CCA for the Scenarios 2a and 2b

Scenario	The analyst considers a composite model	
	Scenario 2a:	Scenario 2b:
Situation	The population functions according to a composite model	The population functions according to a model different from the composite model
	The analyst's model is correct	The analyst's model is incorrect
MCMQ results	Rules for formative measurement models: ○ Convergent validity ✓ Indicator multicollinearity (largest VIF: 1.809 < 3; largest correlation: 0.400 < 0.5) ✓ Size and significance of weights (smallest weight: 0.274) ✓ Size and significance of loadings (smallest loading: 0.504 > 0.5) ○ Predictive validity Model fit criteria: ✓ SRMR = 0.000 ✓ $d_G = 0.000$ ✓ NFI = 1.000	Rules for formative measurement models: ○ Convergent validity ✓ Indicator multicollinearity (largest VIF: 1.809 < 3; largest correlation: 0.400 < 0.5) ✓ Size and significance of weights (smallest weight: 0.148) ✓ Size and significance of loadings (smallest loading: 0.501 > 0.5) ○ Predictive validity Model fit criteria: ✗ SRMR = 0.081 ✗ $d_G = 0.029$ ✗ NFI = 0.896
CCA results	MCMQ confirms measurement quality, and CCA does not reject the model	MCMQ confirms measurement quality, but CCA rejects the model

✓: the criterion indicates no problem; ✗: the criterion indicates a problem; ○: the criterion is not applicable in this scenario

model specification of Scenario 2b (see Table 5). The SRMR and the geodesic distance are both 0 in Scenario 2a and clearly above 0 in Scenario 2b. Similarly, the NFI is 1 in the scenario where the composite model is correctly specified and significantly below 1 in Scenario 2b. Consequently, CCA sounds the alert that the composite model of Scenario 2b is misspecified.

For Scenarios 3a and 3b, the results of CCA and MCMQ are displayed in Table 6. MCMQ's quality criteria for formative measurement models do not indicate any problems: the path coefficient between the two constructs is above 0.7, and thus, following Hair et al. (2020), convergent validity is established (see Table 3). Considering the correlations among the observable variables (see the correlation matrices of Table 2) and the VIF values for the weights of the first emergent variable,¹² multicollinearity is not an issue. Moreover, all weights are sizable and none of the loadings of the first construct are below the suggested threshold of 0.5 (see Table 3).¹⁰

To assess the second emergent variable, MCMQ's rules for reflective measurement models are applied. Again, for both scenarios, no problems are indicated. All loadings associated with the second emergent variable are above the proposed threshold of 0.708 (see Table 3).¹⁰ As no decision criterion for indicator reliability is mentioned by Hair et al. (2020), the squared loadings ranging from 0.680 to 0.755 are regarded as sufficient. Composite reliability is assessed by means of Cronbach's α and Jöreskog's ρ . In both scenarios, the two reliability measures are between the proposed thresholds of 0.70 and 0.95.¹³ Considering the AVE, the values for the second construct are all above the proposed threshold of 0.5.¹⁴ Since there is only one reflective construct, discriminant validity is not assessed. Nomological validity can be regarded as established since the correlation between the construct scores of the second and the first constructs are in line with the researcher's expectations. As a consequence, MCMQ's quality criteria do not indicate any problems with the specified model in Scenario 3b, where the model is indeed misspecified.

The model fit criteria of CCA correctly do not sound alarm in the case of Scenario 3a. The SRMR and the geodesic distance are both zero.¹⁵ Moreover, the NFI is equal to one. However, for Scenario 3b, the SRMR and the geodesic distance are substantially larger than 0. Similarly, the NFI is below 1. As a consequence, CCA correctly alerts a researcher that his or her model is misspecified.

¹² In both scenarios, the VIF values for the weights of the first emergent variable are as follows: w_{11} : 1.264; w_{12} : 1.369; and w_{13} : 1.264.

¹³ Scenario 3a: Cronbach's $\alpha = 0.810$ and Jöreskog's $\rho = 0.888$; Scenario 3b: Cronbach's $\alpha = 0.810$ and Jöreskog's $\rho = 0.886$. It is noted that the scores of emergent variables that are composed of observable variables are fully reliable. However, the two reliability measures produce values below one, as they were originally designed for an underlying common factor model and thus take the assumptions of the reflective measurement model into account.

¹⁴ Scenario 3a: $AVE_{\eta_2} = 0.72$; Scenario 3b: $AVE_{\eta_2} = 0.72$.

¹⁵ If the researcher had specified a reflective measurement model, SEM would have indicated a missfit.

Table 6 Results of MCMQ and CCA for the Scenarios 3a and 3b

Scenario	Scenario 3a: The analyst considers a model that consists of two interrelated emergent variables where one is formed by correlation weights	Scenario 3b: The population functions according to a composite model
Situation	The analyst's model is correct	The analyst's model is incorrect
MCMQ results	<p>Rules for formative measurement models:</p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> Convergent validity ($\beta = 0.800 > 0.70$) <input checked="" type="checkbox"/> Indicator multicollinearity (largest VIF: 1.369 < 3; largest correlation: 0.200 < 0.5) <input checked="" type="checkbox"/> Size and significance of weights (smallest weight: 0.347) <input checked="" type="checkbox"/> Size and significance of loadings (smallest loading: 0.512 > 0.5) <input type="checkbox"/> Predictive validity <p>Rules for reflective measurement models:</p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> Size and significance of loadings (smallest loading: 0.825 > 0.708) <input checked="" type="checkbox"/> Indicator reliability (smallest indicator reliability: 0.680) <input checked="" type="checkbox"/> Composite reliability (Cronbach's $\alpha = 0.810 > 0.7$; CR = 0.888 > 0.7) <input checked="" type="checkbox"/> Average variance extracted (AVE: 0.725 > 0.5) <ul style="list-style-type: none"> <input type="checkbox"/> Discriminant validity <input checked="" type="checkbox"/> Nomological validity (correlation with other construct: 0.800) <input type="checkbox"/> Predictive validity 	<p>Rules for formative measurement models:</p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> Convergent validity ($\beta = 0.908 > 0.70$) <input checked="" type="checkbox"/> Indicator multicollinearity (largest VIF: 1.369 < 3; largest correlation: 0.200 < 0.5) <input checked="" type="checkbox"/> Size and significance of weights (smallest weight: 0.406) <input checked="" type="checkbox"/> Size and significance of loadings (smallest loading: 0.624 > 0.5) <input type="checkbox"/> Predictive validity <p>Rules for reflective measurement models:</p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> Size and significance of loadings (smallest loading: 0.839 > 0.708) <input checked="" type="checkbox"/> Indicator reliability (smallest indicator reliability: 0.704) <input checked="" type="checkbox"/> Composite reliability (Cronbach's $\alpha = 0.810 > 0.7$; CR = 0.886 > 0.7) <input checked="" type="checkbox"/> Average variance extracted (AVE: 0.721 > 0.5) <ul style="list-style-type: none"> <input type="checkbox"/> Discriminant validity <input checked="" type="checkbox"/> Nomological validity (correlation with other construct: 0.908) <input type="checkbox"/> Predictive validity

Table 6 (continued)

Scenario	The analyst considers a model that consists of two interrelated emergent variables where one is formed by correlation weights	
	Scenario 3a:	Scenario 3b:
	The population functions according to a composite model	The population functions according to a model different from the composite model
Situation	The analyst's model is correct	The analyst's model is incorrect
CCA results	Model fit criteria:	Model fit criteria:
	✓ SRMR = 0.000	✗ SRMR = 0.083
	✓ $d_G = 0.000$	✗ $d_G = 0.311$
	✓ NFI = 1.000	✗ NFI = 0.717
Conclusion	MCMQ confirms measurement quality, and CCA does not reject the model	MCMQ confirms measurement quality, but CCA rejects the model
	✓: the criterion indicates no problem; ✗: the criterion indicates a problem; ○: the criterion is not applicable in this scenario	

6 Conclusion and discussion

Dijkstra and Henseler extended SEM and invented a statistical method that assesses the overall fit of a composite model and, making use of the inventor's privilege to name the invention, they called it 'confirmatory composite analysis' (Henseler et al. 2014, see also the author note in that paper). Moreover, Schuberth et al. (2018) fully developed this method and provide evidence for its efficacy. The name of the method was chosen on purpose: The method allows the analysis of composite models; it is confirmatory in nature because it (dis-)confirms whether the specified model is consistent with the data; in the words of Bollen (1989 p. 68): "If a model is consistent with reality, then the data should be consistent with the model". Following the principle that a name should say something about the meaning of the named object, one could thus say that the name 'confirmatory composite analysis' fits the method like a glove. Moreover, the name emphasizes the proximity to CFA: CCA and CFA share everything except for the specified model. Whereas CFA assesses reflective measurement models, also known as common factor models, CCA assesses composite models in which the abstract concept is represented by an emergent variable instead of a latent variable.

In a recent paper published in the *Journal of Business Research*, Hair et al. (2020) took the term 'confirmatory composite analysis' and used it for something else, namely, the evaluation step of PLS-SEM. While it is comprehensible that they employ rebranding to distance their method from negative connotations (evoked, e.g., by papers such as Rönkkö and Evermann 2013; Rönkkö et al. 2015; Dijkstra and Henseler 2015b; Rönkkö et al. 2016; Goodhue et al. 2017), it is unfortunate that they used a name of an extant method that is substantially different. As this paper has shown, Hair et al.'s method neither assesses composite models nor is it suitable for confirmatory purposes. Consequently, it creates unnecessary ambiguity and confusion to call MCMQ 'confirmatory composite analysis', as CCA introduced by Schuberth et al. (2018) was exactly developed for that purpose. Instead, it is recommended either keeping the original name, the 'evaluation step of PLS-SEM', or using the descriptive name 'method of confirming measurement quality', as it was done in this paper.

Whereas CCA has been demonstrated to serve as an effective method to assess composite models [in addition to the scenario analysis presented in this paper, see, for instance, the Monte Carlo simulations conducted by Schuberth et al. (2018) and Schuberth et al. (2020)], so far, no alleged capabilities of MCMQ have been demonstrated. In contrast, this article provided evidence that MCMQ is unable to fulfill its promises. As known in the literature and reconfirmed by this study, MCMQ is not suited to evaluate reflective and causal-formative measurement models. This is mainly due to two reasons: First, model fit assessment plays no role in MCMQ which leaves a researcher uninformed about whether his/her specified model is an acceptable representation of the studied population. Second, several of the evaluation criteria employed in MCMQ assume a reflective measurement model for which PLS-SEM produces inconsistent estimates (e.g., Henseler et al. 2014). Moreover, this study showed that MCMQ fails to disconfirm invalid composite models including models where the emergent variable is

formed by correlation weights (“reflective measurement” in PLS-SEM parlance). This casts additionally doubts about applying evaluation criteria that assume a reflective measurement model, such as Jöreskog’s ρ and AVE, to composite models. From the findings of the scenario analysis, it is concluded that MCMQ lacks sensitivity. While it indeed confirms measurement quality in case of valid measurement, at the same time, it fails to disconfirm measurement quality in cases in which the specified model is in fact invalid. Ergo, the evaluation steps subsumed under MCMQ *confirm* measurement quality, but they do not necessarily assess it.¹⁶

This does not mean that MCMQ’s rules for reflective measurement models should be discarded. In fact, they have been widely proposed in guidelines for the assessment of reflective measurement models estimated by PLS, e.g., Benitez et al. (2020), Müller et al. (2018) and Henseler et al. (2016). However, as shown by this study, they cannot replace the crucial step of model fit assessment in SEM. Every rigorous scientific study should engage in *assessing* the fit of the outer models before confirming their quality.

Apart from promulgating an ineffective method, Hair et al. (2020) make several statements on CCA that are incorrect; not always is the reason of their incorrectness that they confound CCA with MCMQ. Table 7 lists the most important statements and sets the record straight.

Finally, the study at hand provides a very simple rule for researchers who would like to assess the quality of their models: Always assess a model with the method that was developed for this model. Concretely, CFA is used for assessing reflective measurement models (common factor models), SEM is used for assessing causal–formative measurement models, and CCA is used for assessing composite models.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix: Consequences of overestimating the composite reliability by PLS

Traditional PLS as contained in PLS-SEM overestimates composite reliability. This can have detrimental effects on the obtained path coefficients. The following example is based on ideas expressed in the extant literature (in particular, Dijkstra and Henseler 2015b; Goodhue et al. 2017; Henseler 2012).

¹⁶ It is worrying that similar guidelines have been proposed for the model assessment in PLS-SEM, see e.g., Sarstedt et al. (2019), Sarstedt and Cheah (2019), Hair et al. (2019a), Ringle et al. (2020), and Manley et al. (2020). Like MCMQ, these guidelines will also fail to detect many forms of model misspecification.

Table 7 Incorrect statements of Hair et al. (2020) about confirmatory composite analysis and corresponding corrections

Incorrect statement of Hair et al. (2020)	Correction
<p>“The choice of CCA or CFA is deeply intertwined with the choice of PLS-SEM or CB-SEM” (p. 108)</p>	<p>CCA is not bound to any particular estimator. Possible estimators are approaches to generalized canonical correlation analysis (as presented by Schubert et al. 2018), GSCA, and also the iterative PLS algorithm (see for instance Henseler 2015, 2017; Henseler et al. 2016; Henseler and Schubert forthcoming). This is analogous to CFA, for which ML (see Jöreskog 1969), weighted least squares (WLS, see Browne 1984), and also PLS (see Dijkstra and Henseler 2015a, b) can be used as an estimator.</p>
<p>“Henseler et al. (2014) proposed the concept of CCA as a process for confirming measurement models in PLS-SEM” (p. 102); similarly: “CCA is a systematic methodological process for confirming measurement models in PLS-SEM” (p. 104)</p>	<p>Dijkstra and Henseler invented CCA as a comparison of an empirical variance–covariance matrix with the variance–covariance matrix implied by a composite model, for instance, using the SRMR or a bootstrap-based test of exact overall model fit (Henseler et al. 2014). The question is thus: What is part of what? In contrast to what Hair et al. (2020) promulgate, CCA is not part of PLS-SEM, but the iterative PLS algorithm can be a part of CCA (namely, serve as an estimator for the composite model). CCA is thus not “the measurement model assessment steps in PLS-SEM” (Hair et al. 2020, p. 101).</p>
<p>“The statistical objective of CCA ... is confirmation of measurement theory” (p. 102); similarly, “CCA [...] provides a more holistic perspective of measurement properties than EFA [exploratory factor analysis], and it can identify a wider range of measurement concerns when assessing the applicability of measures” (p. 108)</p>	<p>CCA has been developed as a method to model and assess composite models and Schubert et al. (2018) have demonstrated its efficacy. Yet, there is no evidence that CCA is suitable for assessing common factor models, which underlie measurement theory (reflective measurement models). Actually, it is highly questionable whether CCA should be used for analyzing measurement models, considering that there is a special method for this purpose, namely SEM including CFA. Overall, there is no objective reason to use CCA to assess measurement theory.</p>
<p>“CCA can facilitate the assessment of reflective as well as formative measurement models” (p. 108)</p>	<p>Causal–formative measurement models are an extension of reflective measurement models in which causal antecedents are added to the latent variable (c.f. Bollen and Bauldry 2011). To empirically assess measurement theories, researchers could use SEM including CFA. In contrast, CCA is suitable for empirically assessing design theories (Henseler 2017); i.e., CCA assess neither reflective nor (causal-)formative measurement models but composite models. Against the background that the PLS-SEM literature equates the reflective measurement model with a composite model where the emergent variable is made up by correlation weights, it is questionable whether criteria that assume a reflective measurement model are suitable for assessing this particular type of emergent variable.</p>
<p>“Compared to CFA, CCA has several benefits” (p. 108, see also p. 103); similarly, “[i]f researchers are focused on the content validity of constructs, whether reflective or formative measurement models are assumed, then CCA is a superior approach” (p. 108)</p>	<p>Both CFA and CCA consists of the same four steps: model specification, model identification, model estimation, and model assessment. They only differ in the specified model. Whereas the specified model in CFA is a reflective measurement model comprising a latent variable, the specified model in CCA is a composite model comprising emergent variables. Consequently, there is no point in claiming that one <i>method</i> has benefits over the other. However, it may very well be that one <i>model</i> has certain advantages over the other.</p>

Table 7 (continued)

Incorrect statement of Hair et al. (2020)	Correction
<p>“To achieve the objectives of measurement model confirmation objectives in developing or adapting multi-item measures, researchers could use either CFA or CCA” (p. 108)</p>	<p>This statement suggests that CFA and CCA are alternative methods to be used for the same purpose. However, this is wrong. Whereas a CFA assesses a reflective measurement model, a CCA assesses a composite model. The model specification thus dictates whether CFA or CCA be applied. Once the model has been specified, the choice of the method is thus no longer at the discretion of the researcher. Only the estimator can still be determined. The choice of the estimator is based on deliberations with regard to consistency, unbiasedness, robustness, and efficiency.</p>
<p>“One of the most important differences between a CCA and a CFA is the application of so-called goodness-of-fit (GOF) indices” (p. 103); similarly, “goodness of fit measures are not required when executing CCA with formative measurement models” (p. 105), “the execution of CCA within a PLS path model does not require the assessment of fit” (p. 108), and “CCA and PLS-SEM in general should be assessed based on the metrics unique to variance-based SEM, and goodness of fit is not a required metric” (p. 108)</p>	<p>CCA was introduced as a method that assesses the overall fit of a composite model (Henseler et al. 2014; Schubert et al. 2018), i.e., to investigate whether the specified composite model is consistent with the data. Overall model fit assessment is at the core of both CFA and CCA. Conducting CFA and CCA without assessing the model fit makes little sense, because a researcher is left unaware about whether his or her model is wrongly specified. CCA and CFA differ only with regard to the specified model, not with regard to the importance of model fit assessment.</p>
<p>“CCA differs from CFA in that the statistical objective is to maximize variance extracted from the exogenous variables, but in doing so to facilitate prediction of the endogenous constructs, and confirmation of the measurement models” (p. 103); similarly, “CCA should always be considered as a technique when the focus of research is prediction” (p. 108) and “Note that as the final step in CCA the researcher should examine the predictive ability of the structural model.” (p. 106)</p>	<p>None of the original publications on CCA (Henseler et al. 2014; Schubert et al. 2018) mentions prediction as a purpose of CCA. Just like CFA, CCA does not distinguish between exogenous and endogenous constructs. The composite model consists of a set of covarying emergent variables without defining any of them as predictor or predicted variable. Therefore, CCA appears poorly suited for prediction. Finally and as emphasized before, the main goal of CCA is to assess composite models and not to maximize the variance extracted from exogenous variables.</p>

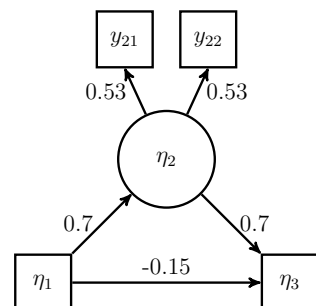
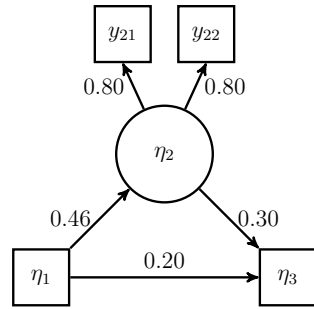
Fig. 3 Population model

Fig. 4 Parameters retrieved by PLS



Let us assume that the population functions according to a model as depicted in Fig. 3.¹⁷ Each of the reflective indicators of the construct η_2 has a reliability of $0.53^2 = 0.28$. The scores of η_2 can then have a reliability of 0.44 at best.

Under these circumstances, PLS will retrieve the parameter values depicted in Fig. 4. The retrieved loadings are severely upward biased, leading to a composite reliability that clearly exceeds 0.7. Based on the presumably satisfactory reliability, a researcher may be confident in the model and interpret its path coefficients, although they are far from true. One path coefficient even has an opposite sign.

References

- Antoninka A, Reich PB, Johnson NC (2011) Seven years of carbon dioxide enrichment, nitrogen fertilization and plant diversity influence arbuscular mycorrhizal fungi in a grassland ecosystem. *New Phytol* 192(1):200–214. <https://doi.org/10.1111/j.1469-8137.2011.03776.x>
- Bagozzi RP (1975) Marketing as exchange. *J Mark* 39(4):32–39. <https://doi.org/10.2307/1250593>
- Barrett P (2007) Structural equation modelling: adjudging model fit. *Pers Individ Differ* 42(5):815–824. <https://doi.org/10.1016/j.paid.2006.09.018>
- Baumgartner H, Homburg C (1996) Applications of structural equation modeling in marketing and consumer research: a review. *Int J Res Mark* 13(2):139–161. [https://doi.org/10.1016/0167-8116\(95\)00038-0](https://doi.org/10.1016/0167-8116(95)00038-0)
- Benitez J, Ray G, Henseler J (2018) Impact of information technology infrastructure flexibility on mergers and acquisitions. *MIS Q* 42(1):25–43. <https://doi.org/10.25300/MISQ/2018/13245>
- Benitez J, Henseler J, Castillo A, Schuberth F (2020) How to perform and report an impactful analysis using partial least squares: guidelines for confirmatory and explanatory IS research. *Inf Manag* 2(57):103168. <https://doi.org/10.1016/j.im.2019.05.003>
- Bentler PM, Bonett DG (1980) Significance tests and goodness of fit in the analysis of covariance structures. *Psychol Bull* 88(3):588–606. <https://doi.org/10.1037/0033-2909.88.3.588>
- Beran R, Srivastava MS (1985) Bootstrap tests and confidence regions for functions of a covariance matrix. *Ann Stat* 13(1):95–115. <https://doi.org/10.1214/aos/1176346579>
- Bollen KA (1989) *Structural equations with latent variables*. Wiley, New York
- Bollen KA, Bauldry S (2011) Three Cs in measurement models: causal indicators, composite indicators, and covariates. *Psychol Methods* 16(3):265–284. <https://doi.org/10.1037/a0024448>
- Borsboom D (2008) Latent variable theory. *Meas Interdiscip Res Perspect* 6(1–2):25–53. <https://doi.org/10.1080/15366360802035497>

¹⁷ To preserve clarity, the error terms are omitted.

- Braojos J, Benitez Je, Llorens J, Ruiz L (2020) Impact of IT integration on the firm's knowledge absorption and desorption. *Inf Manag* 224:103290. <https://doi.org/10.1016/j.im.2020.103290>
- Browne MW (1984) Asymptotically distribution-free methods for the analysis of covariance structures. *Br J Math Stat Psychol* 37(1):62–83. <https://doi.org/10.1111/j.2044-8317.1984.tb00789.x>
- Cepeda Carrión G, Henseler J, Ringle CM, Roldán JL (2016) Prediction-oriented modeling in business research by means of PLS path modeling. *J Bus Res* 69(10):4545–4551. <https://doi.org/10.1016/j.jbusres.2016.03.048>
- Cho G, Choi JY (2020) An empirical comparison of generalized structured component analysis and partial least squares path modeling under variance-based structural equation models. *Behaviormetrika* 47:243–272. <https://doi.org/10.1007/s41237-019-00098-0>
- Cohen P, Cohen J, Teresi J, Marchi M, Velez CN (1990) Problems in the measurement of latent variables in structural equations causal models. *Appl Psychol Meas* 14(2):183–196. <https://doi.org/10.1177/014662169001400207>
- Cole DA, Maxwell SE, Arvey R, Salas E (1993) Multivariate group comparisons of variable systems: MANOVA and structural equation modeling. *Psychol Bull* 114(1):174–184. <https://doi.org/10.1037/0033-2909.114.1.174>
- Diamantopoulos A (2008) Formative indicators: introduction to the special issue. *J Bus Res* 61(12):1201–1202. <https://doi.org/10.1016/j.jbusres.2008.01.008>
- Diamantopoulos A, Riefler P, Roth KP (2008) Advancing formative measurement models. *J Bus Res* 61(12):1203–1218. <https://doi.org/10.1016/j.jbusres.2008.01.009>
- Dijkstra TK (1981) Latent variables in linear stochastic models: reflections on “maximum likelihood” and “partial least squares” methods. PhD thesis, Groningen University
- Dijkstra TK (2017) A perfect match between a model and a mode. In: Latan H, Noonan R (eds) *Partial least squares path modeling*. Springer, Cham
- Dijkstra TK, Henseler J (2011) Linear indices in nonlinear structural equation models: best fitting proper indices and other composites. *Qual Quant* 45(6):1505–1518. <https://doi.org/10.1007/s11335-010-9359-z>
- Dijkstra TK, Henseler J (2015a) Consistent and asymptotically normal PLS estimators for linear structural equations. *Comput Stat Data Anal* 81(1):10–23. <https://doi.org/10.1016/j.csda.2014.07.008>
- Dijkstra TK, Henseler J (2015b) Consistent partial least squares path modeling. *MIS Q* 39(2):297–316. <https://doi.org/10.25300/MISQ/2015/39.2.02>
- Fornell C, Bookstein FL (1982) Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. *J Mark Res* 19(4):440–452. <https://doi.org/10.1177/002224378201900406>
- Fornell C, Larcker DF (1981) Evaluating structural equation models with unobservable variables and measurement error. *J Mark Res* 18(1):39–50. <https://doi.org/10.2307/3151312>
- Glang M (1988) Maximierung der Summe erklärter Varianzen in linear-rekursiven Strukturgleichungsmodellen mit multiplen Indikatoren: Eine Alternative zum Schätzmodus B des Partial-Least-Squares-Verfahrens [maximization of the sum of explained variances in linear-recursive structural equation models with multiple indicators: an alternative to Mode B of the partial least squares approach]. Ph.D. thesis, University of Hamburg, Germany
- Goodhue DL, Lewis W, Thompson RL (2017) A multicollinearity and measurement error statistical blind spot: correcting for excessive false positives in regression and PLS. *MIS Q* 41(3):667–684. <https://doi.org/10.25300/MISQ/2017/41.3.01>
- Grace JB, Bollen KA (2008) Representing general theoretical concepts in structural equation models: the role of composite variables. *Environ Ecol Stat* 15(2):191–213. <https://doi.org/10.1007/s10651-007-0047-7>
- Hair JF, Sarstedt M, Pieper TM, Ringle CM (2012) The use of partial least squares structural equation modeling in strategic management research: a review of past practices and recommendations for future applications. *Long Range Plan* 45(5):320–340. <https://doi.org/10.1016/j.lrp.2012.09.008>
- Hair JF, Hult GTM, Ringle CM, Sarstedt M (2014) *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage, Thousand Oaks
- Hair JF, Hult GTM, Ringle CM, Sarstedt M (2017) *A primer on partial least squares structural equation modeling (PLS-SEM)*, 2nd edn. Sage, Thousand Oaks
- Hair JF, Risher JJ, Sarstedt M, Ringle CM (2019a) When to use and how to report the results of PLS-SEM. *Eur Bus Rev* 31(1):2–24. <https://doi.org/10.1108/EBR-11-2018-0203>
- Hair JF, Sarstedt M, Ringle CM (2019b) Rethinking some of the rethinking of partial least squares. *Eur J Mark* 53(4):566–584. <https://doi.org/10.1108/EJM-10-2018-0665>

- Hair JF, Howard MC, Nitzl C (2020) Assessing measurement model quality in PLS-SEM using confirmatory composite analysis. *J Bus Res* 109(1):101–110. <https://doi.org/10.1016/j.jbusres.2019.11.069>
- Hardin A, Marcoulides GA (2011) A commentary on the use of formative measurement. *Educ Psychol Meas* 71(5):753–764. <https://doi.org/10.1177/0013164411414270>
- Henseler J (2012) Why generalized structured component analysis is not universally preferable to structural equation modeling. *J Acad Mark Sci* 40(3):402–413. <https://doi.org/10.1007/s11747-011-0298-6>
- Henseler J (2015) Is the whole more than the sum of its parts? On the interplay of marketing and design research. University of Twente, Enschede. <http://purl.utwente.nl/publications/95770>. Accessed July 2020
- Henseler J (2016) New developments in PLS path modeling: guest editorial special issue on using partial least squares (PLS) in industrial management. *Ind Manag Data Syst* 116(9):1842–1848. <https://doi.org/10.1108/imds-09-2016-0366>
- Henseler J (2017) Bridging design and behavioral research with variance-based structural equation modeling. *J Advert* 46(1):178–192. <https://doi.org/10.1080/00913367.2017.1281780>
- Henseler J (2018) Partial least squares path modeling: Quo vadis? *Qual Quant* 52(1):1–8. <https://doi.org/10.1007/s11135-018-0689-6>
- Henseler J, Schubert F (forthcoming) Using confirmatory composite analysis to assess emergent variables in business research. *J Bus Res* (**accepted**)
- Henseler J, Ringle CM, Sinkovics RR (2009) The use of partial least squares path modeling in international marketing. In: Sinkovics RR, Ghauri PN (eds) *Advances in international marketing*, vol 20. Bingley, Emerald, pp 277–320. [https://doi.org/10.1108/S1474-7979\(2009\)0000020014](https://doi.org/10.1108/S1474-7979(2009)0000020014)
- Henseler J, Dijkstra TK, Sarstedt M, Ringle CM, Diamantopoulos A, Straub DW, Ketchen DJ, Hair JF, Hult GTM, Calantone RJ (2014) Common beliefs and reality about PLS: comments on Rönkkö and Evermann (2013). *Organ Res Methods* 17(2):182–209. <https://doi.org/10.1177/1094428114526928>
- Henseler J, Ringle CM, Sarstedt M (2015) A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J Acad Mark Sci* 43(1):115–135. <https://doi.org/10.1007/s11747-014-0403-8>
- Henseler J, Hubona G, Ray PA (2016) Using PLS path modeling in new technology research: updated guidelines. *Ind Manag Data Syst* 116(1):2–20. <https://doi.org/10.1108/IMDS-09-2015-0382>
- Hevner AR, March ST, Park J, Ram S (2004) Design science in information systems research. *Manag Inf Syst Q* 28(1):6. <https://doi.org/10.5555/2017212.2017217>
- Hu LT, Bentler PM (1998) Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. *Psychol Methods* 3(4):424–453. <https://doi.org/10.1037/1082-989X.3.4.424>
- Hunt SD (2010) *Marketing theory: foundations, controversy, strategy, and resource-advantage theory*. Taylor & Francis Ltd., London
- Hwang H, Takane Y (2004) Generalized structured component analysis. *Psychometrika* 69(1):81–99. <https://doi.org/10.1007/BF02295841>
- Jöreskog KG (1969) A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika* 34(2):183–202. <https://doi.org/10.1007/bf02289343>
- Jöreskog KG (1971) Statistical analysis of sets of congeneric tests. *Psychometrika* 36(2):109–133. <https://doi.org/10.1007/bf02291393>
- Jöreskog KG, Goldberger AS (1975) Estimation of a model with multiple indicators and multiple causes of a single latent variable. *J Am Stat Assoc* 70(351a):631–639. <https://doi.org/10.1080/01621459.1975.10482485>
- Kettenring JR (1971) Canonical analysis of several sets of variables. *Biometrika* 58(3):433–451. <https://doi.org/10.1093/biomet/58.3.433>
- Kline RB (2015) *Principles and practice of structural equation modeling*. Taylor & Francis Ltd., London
- Kohli AK, Jaworski BJ (1990) Market orientation: the construct, research propositions, and managerial implications. *J Mark* 54(2):1–18. <https://doi.org/10.1177/002224299005400201>
- Lazarsfeld PF (1959) Latent structure analysis. In: Koch S (ed) *Psychology: a study of a science*, vol III. McGraw-Hill, New York, pp 476–543
- Manley SC, Hair JF, Williams RI (2020) Essential new PLS-SEM analysis methods for your entrepreneurship analytical toolbox. *Int Entrep Manag J*. <https://doi.org/10.1007/s11365-020-00687-6>
- McCarthy E (1978) *Basic marketing: a managerial approach*. R.D. Irwin, Homewood
- McDonald RP (1996) Path analysis with composite variables. *Multivar Behav Res* 31(2):239–270. https://doi.org/10.1207/s15327906mbr3102_5

- Mell H, Safra L, Algan Y, Baumard N, Chevallier C (2018) Childhood environmental harshness predicts coordinated health and reproductive strategies: a cross-sectional study of a nationally representative sample from France. *Evol Hum Behav* 39(1):1–8. <https://doi.org/10.1016/j.evolhumbehav.2017.08.006>
- Mulaik SA, James LR, Van Alstine J, Bennett N, Lind S, Stilwell CD (1989) Evaluation of goodness-of-fit indices for structural equation models. *Psychol Bull* 105(3):430–445. <https://doi.org/10.1037/0033-2909.105.3.430>
- Müller T, Schubert F, Henseler J (2018) PLS path modeling—a confirmatory approach to study tourism technology and tourist behavior. *J Hosp Tour Technol* 9(3):249–266. <https://doi.org/10.1108/JHTT-09-2017-0106>
- Nunnally JC, Bernstein IH (1994) *Psychometric theory*, 3rd edn. McGraw-Hill, New York
- Rademaker ME, Schubert F (2020) cSEM: composite-based structural equation modeling. R package version 0.1.0.9000. <https://github.com/M-E-Rademaker/cSEM>. Accessed Mar 2020
- Rasoolimanesh SM, Md Noor S, Schubert F, Jaafar M (2019) Investigating the effects of tourist engagement on satisfaction and loyalty. *Serv Ind J* 39(7–8):559–574. <https://doi.org/10.1080/02642069.2019.1570152>
- Reichenbach H (1956) *The direction of time*. University of California Press, Berkeley
- Reinartz W, Haenlein M, Henseler J (2009) An empirical comparison of the efficacy of covariance-based and variance-based SEM. *Int J Res Mark* 26(4):332–344. <https://doi.org/10.1016/j.ijresmar.2009.08.001>
- Reise SP (1999) Measurement issues viewed through the eyes of IRT. In: Embretson SE, Hershberger SL (eds) *The new rules of measurement: what every psychologist and educator should know*. Psychology Press, New York and London, pp 219–241
- Rhemtulla M, van Bork R, Borsboom D (2020) Worse than measurement error: consequences of inappropriate latent variable measurement models. *Psychol Methods* 25(1):30–45. <https://doi.org/10.1037/met0000220>
- Rigdon EE (2012) Rethinking partial least squares path modeling: in praise of simple methods. *Long Range Plan* 45(5–6):341–358. <https://doi.org/10.1016/j.lrp.2012.09.010>
- Rigdon EE (2016) Choosing PLS path modeling as analytical method in European management research: a realist perspective. *Eur Manag J* 34(6):598–605. <https://doi.org/10.1016/j.emj.2016.05.006>
- Rigdon EE, Sarstedt M, Ringle CM (2017) On comparing results from CB-SEM and PLS-SEM: five perspectives and five recommendations. *Mark ZFP* 39(3):4–16. <https://doi.org/10.15388/0344-1369-2017-3-4>
- Rigdon EE, Becker JM, Sarstedt M (2019) Factor indeterminacy as metrological uncertainty: implications for advancing psychological measurement. *Multivar Behav Res* 53(4):429–443. <https://doi.org/10.1080/00273171.2018.1535420>
- Ringle CM, Sarstedt M, Mitchell R, Gudergan SP (2017) Partial least squares structural equation modeling in HRM research. *Int J Hum Resour Manag* 31(12):1617–1643. <https://doi.org/10.1080/09585192.2017.1416655>
- Rönkkö M, Evermann J (2013) A critical examination of common beliefs about partial least squares path modeling. *Organ Res Methods* 16(3):425–448. <https://doi.org/10.1177/1094428112474693>
- Rönkkö M, McIntosh CN, Antonakis J (2015) On the adoption of partial least squares in psychological research: Caveat emptor. *Pers Individ Differ* 87:76–84. <https://doi.org/10.1016/j.paid.2015.07.019>
- Rönkkö M, McIntosh CN, Antonakis J, Edwards JR (2016) Partial least squares path modeling: time for some serious second thoughts. *J Oper Manag* 47(1):9–27. <https://doi.org/10.1016/j.jom.2016.05.002>
- Sarstedt M, Cheah JH (2019) Partial least squares structural equation modeling using SmartPLS: a software review. *J Mark Analyt* 7(3):196–202. <https://doi.org/10.1057/s41270-019-00058-3>
- Sarstedt M, Hair JF, Cheah JH, Becker JM, Ringle CM (2019) How to specify, estimate, and validate higher-order constructs in PLS-SEM. *Australas Mark J* 27(3):197–211. <https://doi.org/10.1016/j.ausmj.2019.05.003>
- Schubert F, Henseler J, Dijkstra TK (2018) Confirmatory composite analysis. *Front Psychol* 9:2541. <https://doi.org/10.3389/fpsyg.2018.02541>
- Schubert F, Rademaker ME, Henseler J (2020) Estimating and assessing second-order constructs using PLS-PM: the case of composites of composites. *Ind Manag Data Syst*. <https://doi.org/10.1108/IMDS-12-2019-0642>
- Schumacker RE, Lomax RG (2016) *A beginner's guide to structural equation modeling*, 4th edn. Routledge, New York and Milton Park

- Shiau WL, Sarstedt M, Hair JF (2019) Internet research using partial least squares structural equation modeling (PLS-SEM). *Internet Res* 29(3):398–406. <https://doi.org/10.1108/intr-10-2018-0447>
- Shmueli G (2010) To explain or to predict? *Stat Sci* 25(3):289–310. <https://doi.org/10.1214/10-STS330>
- Simon HA (1969) *The sciences of the artificial*. MIT Press, Cambridge
- Steenkamp JBEM, Baumgartner H (2000) On the use of structural equation models for marketing modeling. *Int J Res Mark* 17(2–3):195–202. [https://doi.org/10.1016/S0167-8116\(00\)00016-1](https://doi.org/10.1016/S0167-8116(00)00016-1)
- van Riel AC, Henseler J, Kemény I, Sasovova Z (2017) Estimating hierarchical constructs using consistent partial least squares: the case of second-order composites of common factors. *Ind Manag Data Syst* 117(3):459–477. <https://doi.org/10.1108/IMDS-07-2016-0286>
- Wold HOA (1975) Path models with latent variables: the NIPALS approach. In: Blalock HM (ed) *Sociology quantitative*. Academic Press, Boston, pp 307–357
- Yuan KH (2005) Fit indices versus test statistics. *Multivar Behav Res* 40(1):115–148. https://doi.org/10.1207/s15327906mbr4001_5

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.