Technical Note

A clarification of confirmatory composite analysis (CCA)

Geoffrey S. Hubona, Florian Schuberth, Jörg Henseler

Keywords: CCA, Confirmatory composite analysis, Composite models, Structural equation modeling, Partial least squares structural equation modeling

ABSTRACT

Confirmatory composite analysis (CCA) is a structural equation modeling (SEM) technique that specifies and assesses composite models. In a composite model, the construct emerges as a linear combination of observed variables. CCA was invented by Jörg Henseler and Theo K. Dijkstra in 2014, was subsequently fully elaborated by Schuberth et al. (2018), and was then introduced into business research by Henseler and Schuberth (2020b). Inspired by Hair et al. (2020), a recent article in the International Journal of Information Management (Motamarri et al., 2020) used the same term ‘confirmatory composite analysis’ as a technique for confirming measurement quality in partial least squares structural equation modeling (PLS-SEM) specifically. However, the original CCA (Henseler et al., 2014; Schuberth et al., 2018) and the Hair et al. (2020) technique are very different methods, used for entirely different purposes and objectives. So as to not confuse researchers, we advocate that the later-published Hair et al. (2020) method of confirming measurement quality in PLS-SEM be termed ‘method of confirming measurement quality’ (MCMQ) or ‘partial least squares confirmatory composite analysis’ (PLS-CCA).

We write this research note to clarify the differences between CCA and PLS-CCA.

1. Introduction

Confirmatory composite analysis (CCA) as sketched by Henseler et al. (2014) and elaborated by Schuberth, Henseler, & Dijkstra (2018) is a novel approach to structural equation modeling (SEM). CCA is very similar to confirmatory factor analysis (CFA, Jöreskog, 1969) but instead of assuming a common factor model, CCA assumes a composite model. Consequently, CCA broadens the accessibility and use of SEM to composite models. Like CFA, CCA follows SEM’s typical steps, namely model specification, model identification, model estimation, and model assessment. As is common for SEM, in the last step, namely model assessment, the assessment of the overall model fit plays a crucial role (e.g., Barrett, 2007). In the context of CCA, this means that the discrepancy between the sample variance-covariance matrix and the counterpart implied by the composite model is examined to judge the fit of the specified composite model. Consequently, CCA is a confirmatory approach to (dis-)confirm a researcher’s theory.

The composite model underlying CCA was formally introduced by Dijkstra (2017). In contrast to the reflective and (causal-)formative measurement model known from SEM (Bollen & Bauldry, 2011), in the composite model, abstract concepts are represented as emergent variables, i.e., composites, which are linear combinations of other variables. Typically the covariances among the variables forming an emergent variable are not constrained, i.e., these variables can freely covary. However, the composite model puts constraints on the covariances between the variables forming an emergent variable and the other variables in the model not forming this emergent variable, namely, that these covariances are proportional. Specifically, the composite model assumes that all information between the variables forming the emergent variable and the other variables not forming this emergent variable is solely conveyed by the emergent variable. The composite model appears to be a natural choice to model formed concepts, i.e., abstract concepts that are emergent and artificial and are assumed to be formed by their ingredients (Benitez et al., 2020; Henseler, 2015, 2017, 2019).
Schuberth (2020) previously noted that Hair et al. (2020) research note/technical note, giving a detailed overview of CCA along with existing, well-known estimation steps from partial least squares structural equation modeling (PLS-SEM, Hair et al., 2011). However, the authors also chose to label their technique as ‘confirmatory composite analysis’ citing Hair et al. (2020). What Hair et al. (2020) describe as ‘confirmatory composite analysis’ summarizes the many long-standing approaches and techniques used to evaluate the quality of PLS-SEM measurement models. These models comprise reflective or formative PLS-SEM measurement models to operationalize concepts such as attitudes and beliefs, which cannot be directly observed but which are typically assumed to exist in nature. It is noted that reflective and formative PLS-SEM measurement models differ substantially from the reflective and (causal-formative) measurement models known from the SEM literature (Rigdon, 2012), see also Section 5. In the remainder of this paper, we label the formative measurement models on which CCA is based as CCA-PLS ‘PLS-CCA’ so as not to confuse it with the original CCA.

Summarizing CCA and PLS-CCA under one name is unfortunate because the two have different purposes, mathematical foundations, and objectives. According to Hair et al. (2020) p. 104, PLS-CCA “is a systematic methodological process for confirming measurement models in PLS-SEM.” Yet, according to Schuberth et al. (2018), and in contrast to PLS-CCA, CCA was not developed to assess measurement models, whether reflective or formative. Instead, it was proposed to specify and assess composite models. Schuberth (2021b) discusses additional fundamental differences between PLS-CCA and CCA, noting that the PLS-CCA “does not require the assessment of fit” (Hair et al., 2020, p. 108). In contrast, assessing overall model fit is essential to completing a CCA. Furthermore, while PLS-CCA is inextricably linked to PLS-SEM, CCA is not. Although CCA originated in the PLS-SEM framework and the iterative partial least squares (PLS, Wold, 1975) algorithm can be used to estimate weights, other estimators such as Kettingen’s approaches to generalized canonical correlation analysis (GCCA, Kettinging, 1971) or maximum likelihood can be applied as well (Henseler & Schuberth, 2021a; Schuberth, 2021a; Schuberth et al., 2018). Thus, CCA is clearly not tied to the iterative PLS algorithm and PLS-SEM.

Against the background of the situation described above, the purpose of this paper is to clarify the distinction between CCA (Henseler et al., 2014; Schuberth et al., 2018) and PLS-CCA (Hair et al., 2020). Hereby, we respond to a specific request from the editor of the International Journal of Information Management to write a short opinion paper/research note/technical note, giving a detailed overview of CCA along with a review of articles that have already utilized CCA. Henseler & Schuberth (2020) previously noted that Hair et al. (2020) “mistook CCA for the measurement model evaluation step of partial least squares structural equation modeling” (p. 1) and provide a detailed description of the four steps of CCA including the mathematical foundations and an empirical demonstration of CCA. Although our paper is similar to the ones of Henseler and Schuberth (2020) and Schuberth (2021b), it shows its own merits. Specifically, in this paper, we focus on the differences between CCA and PLS-CCA and offer conclusions highlighting the differing purposes and procedures of CCA and PLS-CCA.

The remainder of the paper is structured as follows. Section 2 provides an overview of CCA including a review of articles that have already utilized CCA. To properly compare CCA and PLS-CCA, in Section 3, we introduce CCA as an approach to SEM consisting of the following four steps: (1) specifying the model; (2) ensuring that the model is identified; (3) estimating the model parameters; and (4) assessing the model, which are elaborated in the following subsections.

Similarly to CCA also the composite model on which CCA is based emerged in the context of PLS-SEM. Originally, the iterative PLS algorithm was developed against the backdrop of a latent variable model, for which it is known to produce inconsistent estimates (e.g., Dijkstra, 1985). To address this problem, Dijkstra (2017) formally introduced the composite model to provide a model specification that can be consistently estimated by the iterative PLS algorithm using Mode B for calculating the weights. As shown in Dijkstra (2017), the composite model can be detached from PLS, and other estimators can be used for its estimation.

2. Overview of CCA

Historically, CCA emerged from PLS-SEM in 2014 as a response to the critiques raised by Rönkkö and Evermann (2013) about PLS-SEM, particularly that no evaluation criteria exist that reliably distinguish between correctly and wrongly specified models (Henseler et al., 2014). To address this concern, CCA was proposed as the step of overall model fit assessment: “The model-implied covariance-matrix of the PLS path model is computed using Equation 2, and we determine the exact fit of the composite factor model by means of bootstrapping the conventional likelihood function. In essence, this constitutes a confirmatory composite analysis.” (Henseler et al., 2014, p. 194). Although ten authors were involved in the study of Henseler et al. (2014), the authors’ note clearly states that: “[t]he formal development of the composite factor model is attributable to the second author [i.e., Theo K. Dijkstra], and in collaboration with the first author [i.e., Jörg Henseler] developed into the concept of confirmatory composite analysis.” Subsequently, in 2018 Florian Schuberth published together with his co-authors Jörg Henseler and Theo K. Dijkstra the first full elaboration of CCA, see Schuberth et al. (2018). CCA can be regarded as a special case of GCCA and as a generalization of extended redundancy analysis (Hwang, 2005).

In their article, Schuberth et al. (2018) introduce CCA as an approach to SEM consisting of the following four steps: (1) specifying the model; (2) ensuring that the model is identified; (3) estimating the model parameters; and (4) assessing the model, which are elaborated in the following subsections.

2.1. CCA Step 1: specifying the model

Typically, in explanatory statistical modeling, there is a theory that is believed to explain some part of the world. In the context of SEM, this means that we have a theoretical model which links abstract concepts to each other. To statistically assess our theory, as a first step, we have to convert the theoretical model into a statistical model. However, the manner in which we model a concept depends on the researcher’s understanding about the concept. Current literature argues that there are at least two types of concepts (Schuberth et al., 2018; Benitez et al., 2020a; Henseler, 2015, 2017; Henseler & Schuberth, 2021b, 2021).

---

1 It is noted that in the literature, this type of concept is also labeled as forged concept (Henseler, 2021) and artifact (Henseler, 2017). We use the term ‘formed concept’ to emphasize the nature of this type of concept. Moreover, while Schuberth et al. (2018) and Henseler & Schuberth (2020) argue for the use of the composite model to operationalize formed concepts, recently it has also been suggested to employ the composite model to model behavioral concepts (see, e.g., Rigdon, Sarstedt, & Ringle, 2017). This notion assumes that both latent and emergent variables serve as a proxy for behavioral concepts (Sarstedt, Hair, Ringle, Thiele, & Gudergan, 2016).

2 We thank an anonymous reviewer for laying this connection.
There are behavioral concepts (see Fig. 1) that are assumed to exist in nature but which cannot be directly observed such as attitudes and opinions. Because these theoretical concepts cannot be observed directly, we collect measures of these concepts, for example by means of responses to survey questions. Hence, there is an assumed causal relationship between the concept and its observed variables. To model this statistically, the reflective measurement model, which is also called the common factor model, is typically employed. In the reflective measurement model, the concept is represented by a latent variable in the statistical model. The dominant statistical approach to assess reflective measurement models is CFA (Jöreskog, 1969).

However, following the reasoning of Henseler & Schuberth (2020), there are also contrived or formed concepts (see Fig. 1). Examples of such formed concepts might include capabilities, aggregated indices, and overall values. For more examples of formed concepts, see Subsection 2.6 about empirical studies that have utilized CCA. Formed concepts are assumed to not exist in nature per se, but which are constructed or ‘designed’. Against this background, the reflective measurement model does not appear to be a suitable way of operationalizing formed concepts. As an alternative, the composite model can be used to operationalize these concepts (Benitez et al., 2020a; Henseler, 2017; Henseler & Schuberth, 2020; Schuberth et al., 2018). In the original composite model, the concept which is ‘formed’ (or constructed or designed) is represented as an emergent variable that is a composite of observed variables. However, emergent variables of latent and/or emergent variables are also conceivable (e.g., van Riel, Henseler, Kemény, & Saso-voy, 2017; Schuberth, Rademaker, & Henseler, 2020). In the composite model, the relationships from the observed variables to the concept are assumed to be a definitorial one rather than a causal one. The effects of the observed variables as measured by their weights are assumed to define (rather than ‘cause’) the concept. Moreover, in the original composite model, the variables making up the emergent variables are assumed to be free from random measurement error. Finally, because formed concepts are assumed to emerge within their environment and are therefore context-specific, variables that relate to the formed concepts, next to those variables making up the formed concept, need to be specified.

2.2. CCA Step 2: identifying the model

In Step 2 of CCA, a researcher needs to ensure that the specified model is identified. Model identification means that we can retrieve a unique set of the model parameters from the variance-covariance matrix of the observed variables. As shown in Dijkstra (2017), model identification is of the same importance for composite models as for latent variable models. To achieve model identification in CCA, we must fix the scale of the emergent variables. This can be accomplished, for example, by scaling the weights to ensure that each emergent variable has a unit variance (which is automatically done if the PLS algorithm is applied to estimate the model parameters). Moreover, no emergent variable is allowed to be isolated in the model. For an elaboration of the identification rules for composite models, we refer to Dijkstra (2017) and Henseler & Schuberth (2020).

2.3. CCA Step 3: estimating the model

Once the model is identified, we proceed to Step 3 of CCA, model estimation. Typically, the correlations between the variables forming an emergent variable, the weights, and thus the correlations between the emergent variables and the other variables in the model are unknown. Hence, in empirical research, we have to estimate those parameters and we must choose an estimator for this purpose. Preferably, the researcher uses a consistent estimator, i.e., its estimates converge in probability towards the population parameters. In CCA, we can use the iterative PLS algorithm to consistently estimate the parameters of the composite model (Dijkstra, 2017). However, other consistent estimators can also be used, including Kettenring’s (1971) approaches to GCA such as MAXVAR or MINVAR (Dijkstra, 2017), generalized structured component analysis (GSCA, Hwang and Takane, 2004; Hair et al., 2017; Cho & Choi, 2020), and maximum likelihood (if properly designed for the composite model, see Henseler & Schuberth (2021a), and Schuberth (2021a)). For a detailed explanation of the various estimators and their properties for the composite model, the interested reader is referred to the cited sources.

2.4. CCA Step 4: assessing the model

After the model has been estimated, we can proceed with Step 4, the assessment of the model. In CCA, model assessment is accomplished in two steps, by: (1) assessing the overall fit of the specified model; and (2) assessing the emergent variables themselves and their corresponding parameter estimates. Overall model fit is important to CCA to gauge whether our model corresponds to ‘reality.’ That is, can we falsify our underlying theory? As in SEM, overall model fit can be assessed by testing the exact overall model fit and by utilizing fit indices. Both kinds of overall model fit assessment are based on the discrepancy between the
estimated model-implied and the empirical variance-covariance matrix of the observed variables. To test the overall model fit in CCA, originally a bootstrap-based test has been proposed (Beran & Srivastava, 1985; Schuberth et al., 2018). In doing so, bootstrapping in combination with discrepancy measures including the standardized root mean squared residual (SRMR), the geodesic distance and the squared Euclidean distance can be used to test the null hypothesis of exact model fit. To do so, we test whether the model-implied variance-covariance matrix based on the population parameters equals the population variance-covariance matrix of the observed variables \( \mathbf{H}_0: \mathbf{\Omega}(\theta) = \mathbf{\Omega} \). If the discrepancy between the estimated model-implied and the sample variance-covariance matrices of the observed variables is significantly different from zero, we have empirical evidence that our model is not a proper description or representation of the population, and thus we have evidence contrary to the proposed theory (falsification). We can also use fit indices to assess the approximate model fit. Although they have been originally developed for the latent variable model, they have recently been proposed to evaluate composite models (Schuberth, Rademaker, & Henseler, 2021). It is important to note that the variance-covariance matrix implied by the composite model has to be applied. Potential candidates are the SRMR (Bentler, 1995; Henseler et al., 2014) and the Goodness-of-fit index (GFI, Jöreskog and Sörbom, 1989; Cho, Hwang, Sarstedt, & Ringle, 2020). These measures quantify the degree of model (mis-)fit, but they are descriptive and not inferential. The use of these measures typically entails cutoff values to judge the fit of the model. Consequently, this practice has been criticized as subjective and arbitrary (Marsh, Hau, & Wen, 2004).

If the model fit is regarded as acceptable, then the estimated model parameters should be inspected to see if they align with the underlying theory. Are the values of the estimated model parameters within expected and acceptable ranges? Are they in the expected direction (i.e., positive or negative)? Are the estimated model parameters statistically significant? Potential multicollinearity issues for the weights should be examined, particularly if the weights are not statistically significant or show a directionally-unexpected sign. To assess for multicollinearity, statistical measures such as the variance inflation factor or the tolerance can be used.

2.5. Software to conduct CCA

Researchers can draw from various software packages to conduct a CCA. The choice for certain software is mainly driven by the choice about the estimator used in CCA. If the iterative PLS algorithm is the estimator of choice, commercial software such as ADANCO (Henseler & Dijkstra, 2015) or open-source R packages such as cSEM (Rademaker & Schuberth, 2021) can be used. For instance, Fig. 2 depicts a CCA model specification in ADANCO. In the case that researchers want to employ GSCA as an estimator in CCA, they can use GSCAPro (Hwang et al., 2021), or the open-source R packages gscap (Hwang et al., 2017) and cSEM. It should be noted that currently neither GSCAPro nor gscap allow testing of the exact overall model fit. To apply Kettenring’s approaches to GCCA in CCA, the open-source R package cSEM provides several implementations. Finally, if maximum likelihood is the estimator of choice, researchers can draw from commercial software such as Mplus (Muthén & Muthén, 1998) or the open-source R package lavaan (Rosseel, 2012). It is noted that most software implementing the iterative PLS algorithm or GSCA require the specification of a structural model. In this case, a saturated structural model can be specified to mimic the situation in which all constructs are freely correlated. For guidelines on conducting CCA using the iterative PLS algorithm as implemented in ADANCO and cSEM, the interested reader is referred to Henseler and Schuberth (2021a). Similarly, Henseler and Schuberth (2021a) and Schuberth (2021a) present a model specification that allows for estimating CCA with Mplus and lavaan.

2.6. Review of articles that have utilized CCA

Although CCA is a very new method, it has already found widespread application across various disciplines including family business (Pittino, Martínez, Chirico, & Galvan, 2018; Ruiz-Palomo, Dieguez-Soto, Duréndez, & Santos, 2019), information systems research (Benitez et al., 2018; Brajos et al., 2019; Brajos et al., 2020; Castillo et al., 2021; Cheng et al., 2021; Syed et al., 2020), decision sciences (Benitez et al., 2020b; Felipe, Leidner, Roldan, & Leal-Rodriguez, 2020; Lin, Luo, & Benitez, 2020), innovation management (Cegarra-Navarro, Papa, Garcia-Perez, & Flano, 2019; Cegarra-Navarro, Ruiz, Martinez-Caro, & Garcia-Perez, 2021), knowledge management (Martelo-Landegro, Cegarra-Navarro, & Cepeda-Carrion, 2019; Roldan, Real, & Sanchez Ceballos, 2018; Sanchez-Polo, Cegarra-Navarro, Cillo, & Wensley, 2019), management education (Hernandez-Perlines et al., 2016; Rueda et al., 2017), marketing (Foltean, Trif, & Tuleu, 2019), organizational behavior (Felipe, Roldan, & Leal-Rodriguez, 2016), project management (Benitez-Avila, Hartmann, Dewulf, & Henseler, 2018), service management (Yiu, Ngai, & Lei, 2020), supply chain management (Wei, Ke, Liu, & Wei, 2020), environmental research and public health (Sánchez-Hernández, Stankevičiūtė, Robina-Ramírez, & Díaz-Caro, 2020) and tourism management (Rasoolimanesh, Md Noor, Schuberth, & Jaafar, 2019; Sanchez-Franco, Cepeda-Carrion, & Roldán, 2019).

3. Overview of PLS-CCA

To clarify the differences between CCA and PLS-CCA, in the following we present PLS-CCA. Hair et al. (2020) describe PLS-CCA as “a series of steps executed with PLS-SEM to confirm both reflective and formative measurement models” (p. 1). It is important to notice that when Hair et al. speak of reflective measurement models, they refer to PLS Mode A; and when they speak of formative measurement models, they refer to PLS Mode B (see, e.g., the glossary in Hair et al. (2014). The series of steps inclusively consist of seven steps to confirm reflective PLS-SEM measurement models, and five steps to confirm formative PLS-SEM measurement models. Subsections 3.1 and 3.2 present the steps involved to assess reflective and formative PLS-SEM measurement models.

3.1. PLS-CCA with reflective measurement models

According to Hair et al. (2020) the seven steps listed below should be followed to execute PLS-CCA with reflective PLS-SEM measurement models:

Step 1: Assess the indicator loadings and their significance.
Step 2: Square the individual indicator loadings to provide a measure of the amount of variance shared between the individual indicator variable and its associated construct.
Step 3: Measure the reliability of the construct using either Cronbach’s alpha or the composite reliability.
Step 4: Measure convergent validity using Average Variance Extracted (AVE).
Step 5: Use discriminant validity to measure the distinctiveness of the construct.
Step 6: Use nomological validity as an additional measure to assess construct validity.
Step 7: Assess predictive validity as the extent to which a construct score predicts scores on some criterion measure.

3.2. PLS-CCA with formative measurement models

According to Hair et al. (2020) the five steps listed below should be followed to execute PLS-CCA with formative PLS-SEM measurement models:

Step 1: Assess convergent validity as the extent to which the formative construct is positively correlated with a reflective measure(s)
with respect to the role of model identification and in their intrinsic relationship with the PLS method specifically. While model identification plays an important role in CCA, it is not directly relevant to the measurement model evaluation techniques used in PLS-CCA. Furthermore, the PLS-CCA techniques can be positioned entirely within the evaluation step of PLS-SEM and therefore PLS-CCA is inextricably linked to the PLS-SEM framework. On the other hand, for CCA, various estimators can be employed to obtain the model parameters, such as Kullback-Leibler’s (1971) approaches to GCCA (Dijkstra, 2017) or maximum likelihood (Henseler & Schuberth, 2021a; Schuberth, 2021a). Moreover, the estimates for CCA can be derived using the iterative PLS algorithm, but this is not mandatory. Hence, CCA is by no means tied to the PLS-SEM framework.

CCA and PLS-CCA also differ fundamentally on the assessment of model fit. Similar to other forms of SEM, the assessment of overall model fit is a critical consideration with CCA. In brief, CCA relies on the discrepancy between the estimated model-implied and the empirical variance-covariance matrix of observed variables to assess overall model fit. In contrast, PLS-CCA and PLS-SEM contribute less value to overall model fit assessment and foster doubt about their usefulness in this regard (see, e.g., Hair et al., 2020; Hair et al., 2019). In fact, PLS-CCA excludes model fit as a criterion for PLS-SEM model validation.

The two techniques also differ on their reputed suitability for predictive research. CCA as originally proposed does not include a step for the assessment of the predictive power of a model. In contrast, PLS-CCA explicitly evaluates the model’s predictive power by means of PLS-predict (Hair et al., 2020; Shmueli et al., 2016). This might be explained by the two approaches’ differing natures. PLS-SEM, and thus PLS-CCA, is mainly used for causal-predictive modeling (see e.g., Hair et al., 2019; Chin et al., 2020), while CCA is a confirmatory approach that aims at modeling the underlying data generating process from which the sample at hand was drawn.

5. Critical appraisal of CCA and PLS-CCA

CCA is a confirmatory technique in analyzing composite models by its basic nature. Therefore, the term “Confirmatory Composite Analysis” is the appropriate descriptive label for CCA. In contrast, it is questionable whether PLS-CCA is best described as “confirmatory” since it

---

3 In empirical research, models containing a mixture of latent and emergent variables are expected. In such a situation one could speak of confirmatory composite/factor analysis (CCFA). In this context, approaches should be used that can cope with both latent and emergent variables such as PLSc (Dijkstra & Henseler, 2015a, 2015b), integrated generalized structured component analysis (IGSCA, Hwang et al., 2020) or maximum likelihood estimation (Schuberth, 2021a).

4 For a discussion about the importance of overall model fit assessment, the interested reader is referred to the special issue in Personality and Individual Differences (Vernon & Eysenck, 2007).
ignores key concepts of confirmatory research, such as overall model fit assessment. The objective of PLS-CCA is to assess the quality of reflective and formative PLS-SEM measurement models. Consequently, one might argue that the term “confirmatory” does not align well with the nature of the PLS-CCA technique and guidelines.

The reflective and formative measurement models in PLS-SEM, and thus in PLS-CCA, differ from the common understanding of reflective and (causal-)formative measurement in the SEM literature. While in the SEM literature reflective and (causal-)formative measurement models contain a latent variable representing the abstract concept under investigation (Bollen & Bauldry, 2011), the reflective and formative PLS-SEM measurement models are both composite models. Specifically, in the reflective PLS-SEM measurement model the emergent variable is assumed to be formed by correlation weights, i.e., PLS Mode A, while in the formative PLS-SEM measurement, the composite is assumed to be formed by regression weights, i.e., PLS Mode B (Rigdon, 2012). In fact, the latter is the same composite model as in CCA, i.e., a model where the emergent variable is assumed to be composed by other variables and all the information between these variables and other variables in the model is solely conveyed by the emergent variable. This also explains why most of the PLS-CCA evaluation steps for formative measurement models, such as assessing the size and significance of the weights and composite loadings, have also been proposed in model assessment step of CCA (Henseler & Schuberth, 2020, 2021a). In contrast, the reflective PLS-SEM measurement model is a special case of this composite model. It additionally assumes that the emergent variable explains as much as possible of the variance of the variables forming the emergent variable (Cho & Choi, 2020). The benefits of this additional assumption for empirical research and in particular theory modeling still needs to be explored. Since this special type of composite model is nested in the general composite model, both composite models can be consistently estimated by PLS using Mode B. In contrast, PLS using Mode A will likely lead to inconsistent parameter estimates for the general composite model (Dijkstra, 2017).

In terms of methodological rigor, CCA is a statistical method that is characterized by the same statistical rigor as other forms of SEM such as CFA. It has been empirically demonstrated that CCA is able to discriminate between correctly specified and misspecified composite models (Schuberth et al., 2018, 2020; Schuberth, 2021b). However, like other inferential techniques, the statistical power of the tests applied in CCA depends on the sample size. In contrast, while the guidelines embedded in PLS-CCA might be useful to assess the adequacy of PLS-SEM measurement models, they do not have a central methodological focus outside of this context. In fact, the evidence regarding the efficacy of PLS-CCA’s evaluation steps has been questioned in the literature (McIntosh, Edwards, & Antonakis, 2014; Ronkkö, McIntosh, Antonakis & Edwards, 2016; Schuberth, 2021b). This is due to the fact that most of the metrics employed to validate reflective PLS-SEM measurement models such as indicator reliability, Cronbach’s alpha, composite reliability, average variance extracted, and the heterotrait-monotrait ratio of correlations have been derived under the common factor model, i.e., a reflective measurement model known from SEM comprising a latent variable. Interpreting these metrics for composite models, regardless of whether the emergent variable is composed by regression or correlation weights is questionable, particularly, if inconsistent PLS-SEM estimates are used for their calculation.

PLS-SEM, and thus PLS-CCA, is regarded as a causal-predictive approach while CCA is a confirmatory approach. Consequently, in CCA model fit assessment plays a crucial role whereas in PLS-CCA the step of model fit assessment is omitted and instead predictive measures are considered. In fact, “[PLS–]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.” (Hair et al., 2020, p. 108). This is unfortunate because researchers miss an important opportunity to identify misspecified models (see, e.g., Schuberth, 2021b) and highlighted by the fact that ‘the wrong’ model can sometimes predict better than the correct one” (Shmueli, 2010). Moreover, “most IS researchers do not study research questions where predictive modeling would be applicable, but focus on theory-testing that requires explanatory models.” (Evermann & Ronkkö, 2021). This additionally raises concerns about the benefits of causal-predictive modeling for information systems and management research.

6. Conclusion

The purpose of CCA is to model and assess composite models. In contrast, the purpose of PLS-CCA is to confirm the quality of reflective and formative PLS-SEM measurement models. Furthermore, while CCA is not tied to PLS, the latter can be used as an estimator for CCA in some circumstances. In contrast, PLS-CCA is tied to PLS-SEM. It is the measurement model confirmation step of PLS-SEM. Whereas CCA entails a structured approach for model specification, model identification, model estimation, and model assessment, PLS-CCA does not. The PLS-CCA heuristics and guidelines apply to the quality confirmation of PLS-SEM measurement models. In this regard, PLS-CCA prescribes seven steps to assess reflective PLS-SEM measurement models and five steps to assess formative PLS-SEM measurement models. Whereas the assessment of overall model fit is central to CCA, PLS-CCA does not require the assessment of model fit. Finally, while there exists both mathematical and empirical evidence to support CCA, there is counterevidence of the efficacy of PLS-CCA. Consequently, we recommend that future research studies pay direct attention to the explicit differences between PLS-CCA and CCA and use the appropriate term in the proper context.

Author statement

Prof. Jörg Henseler acknowledges a financial interest in ADANCO and its distributor, Composite Modeling. Besides this, we have no competing interests to declare.

CRediT authorship contribution statement

Geoffrey Hubona: Writing – original draft, Writing – review & editing; Florian Schuberth: Writing – review & editing, Supervision; Jörg Henseler: Conceptualization, Writing – review & editing.

Declarations of interest

Prof. Jörg Henseler acknowledges a financial interest in ADANCO and its distributor, Composite Modeling.

References


