Creative Business and Sustainability Journal (CBSJ)
Vol.44 No.2 July – December 2022, pp.22-40

Using Latent Variables for Confirmatory Composite Analysis

Chanta Jhantasana*

Abstract

In general, all constructs in a confirmatory factor analysis (CFA) are latent variables. Should all constructs also be emergent variables, a hypothetical construct stemming from a latent variable that has received little attention in studies, then a confirmatory composite analysis (CCA) is a possibility. This study employed latent variables as emergent variables in order to conduct a CCA. The latent variables were related to an individual's traits, attitudes, or behavioral notions, such as satisfaction, trust, or loyalty. An emergent variable is composed of data on capabilities, values, indices, therapies, and interventions, as well as an artifact or design idea. CFA was used to analyze the satisfaction, trust, and loyalty of 200 Lazada and Shopee customers, and additional emergent variables were created from latent variables for the CCA. The study demonstrates that emergent variables can arise from latent variables and that CCA is more accurate than CFA.

Keywords: Partial Least Squares Path Modeling; Confirmatory Factor Analysis; Confirmatory Composite Analysis

Received: March 30, 2022 | **Revised**: May 20, 2022 | **Accepted**: July 1, 2022

* Associate Professor, Management Science, Valaya Alongkorn Rajabhat University Under the Royal Patronage.

Introduction

Structural equation modelling (SEM) is used in several disciplines to test empirical theories, making it a desirable method (Schuberth et al., 2020). There are three types of SEMs: The structural model of the Cowles Commission, the covariance-based structural equation model (CB-SEM) of Karl Gustav Jöreskog, and the PLS-SEM of Hermann Wold. The first is more widely used in the econometric group, while Jöreskog and Wold's structural equation models are more commonly used in social sciences. The Jöreskog structural equation model uses the CB-SEM maximum likelihood estimation, which Jöreskog (1969) proposed for CFA. The aim of CFA is to determine if given data suits a proposed measurement model. The CB-SEM is more stable but more restrictive, it includes normal distribution, and uses latent variables. It provides a reflection of the relationship between the indicators and the latent variables. The CB-SEM should be more adaptable since it is hardly convergent. Wold's SEM uses an ordinary least squares estimate based on a variance model called PLS-SEM or PLS-PM (path modeling). It measures the relationship between the indicators and the composite variables as the reflective measures, and uses the causal-formative variables as the formative. Rönkkö and Everman (2013) investigated the problem of composite use reflectively, including causal-formative use in a formative model (Cadogan, & Lee, 2013). As a result, Dijkstra and Henseler (2015a,b) developed consistent partial least squares (PLSc) as a factor or latent variable for a reflective model. Henseler et al. (2014), Henseler (2017b) and Schuberth et al. (2018) developed a CCA using composite or emergent variables for a formative model. Henseler (2017a) states that if all the construct variables in the model are latent, the structural equation model is CFA. Alternatively, if all the construct variables in the model are emergent, then the CCA is available. PLS-SEM can now measure both the CFA and the CCA, and can obtain a comfortable convergence advantage, with no normal distribution requirement, while the CB-SEM can only measure the CFA. Both the CFA and CCA have four stages: Specifying the model, identifying the model, estimating the model, and assessing the model (Henseler, 2017a; Henseler & Schuberth, 2020; Hubona et al., 2021; Muller et al., 2018; Schuberth et al., 2018). The software packages most often used for PLS-SEM are ADANCO, GSCA, SmartPLS, and WrapPLS, all of which allow researchers to develop the model in the first stage and acquire results in the final stage as well as to assess the model using those tools. The PLS-SEM, or variance-based SEM, needs no normalized data and does not suffer from identification or convergence problems (Arnett, et al., 2003; Reinartz et al., 2004).

In general, CFAs and CCAs are distinct from each other in that the CFA utilizes behavioral data such as traits and attitudes, whereas the CCA does not rather, it uses design ideas that incorporate data on capabilities, values, indices, therapies, and interventions (Schubert et al., 2018; Hubona et al., 2021). The phrase "emergent variables" refers to variables that have different meanings in different projects based on the design. Additionally, Hubona et al. (2021, p.3) state that "However, emergent variables of latent and/or emergent variables are also conceivable". Thus, this research adopts this idea of generating emergent variables from latent variables of customer satisfaction, trust, and loyalty for the online shopping platforms Lazada and Shopee. Except for Jhantasana (2022), there are no studies to our knowledge that construct emergent variables from latent variables. CFA has more restrictions than CCA in that it can only use latent variables, while CCA can use all variables, including latent, observable, and emergent.

Thus, this study used latent variables for customer satisfaction, trust, and loyalty towards Lazada and Shopee to undertake CFA. Additionally, this study proposed to construct emergent variables from latent variables (customer satisfaction, trust, and loyalty) in order to construct a CCA. As a result, the goal of this study was to compare the outcomes of CFA and CCA in terms of satisfaction and loyalty for Lazada and Shopee, as well as the mediating influence of trust.

Literature Review

Customer satisfaction, trust, and loyalty towards Lazada and Shopee's online shopping platforms

A student is an excellent illustration of a modern client who makes digital or online purchases, the Lazada and Shopee consumer shopping platforms, which are the most popular online shopping sites in Thailand, Malaysia, Singapore, Indonesia, Vietnam, and the Philippines (Tran, 2019), were used in this investigation for student satisfaction, trust, and loyalty. The indirect effect of satisfaction on loyalty was examined, as was the mediating role of trust. Further, satisfaction served as a predictor of both trust and loyalty (Horppu et al., 2008; Yoon, 2007; Yoon, 2002). According to multiple studies, trust serves as a mediating variable in a model of consumer-brand interaction that encompasses several aspects (Chaudhuri & Holbrook, 2001) including loyalty (Guenzi et al. 2009). In this research, trust is used as a mediating factor to assess consumer satisfaction and loyalty of those using the Lazada and Shopee applications. The indirect impact of this research may be used to validate the efficiency of a model since if it is inaccurate or has problems, such as small sample size, it will be unable to demonstrate the mediation effect (Nitzl et al, 2016).

Partial least squares path model

The PLS-PM is used in many areas such as marketing (Hair et al., 2012), human resources (Richter et al., 2016; Ringle et al., 2018), accounting (Nitzl, 2016), and strategic management (Hair et al., 2012). The PLS-SEM is a composite variable model, as compared to the CB-SEM with a variable factor (Rigdon et al., 2017). Each uses a different algorithm for latent variables (Richter et al., 2016). The PLS-SEM needs to maximize the variance of endogenous constructs, while the CB-SEM minimizes the discrepancy between model covariance metrics and data covariance metrics (Rigdon et al., 2017). The PLS-SEM's advantage is predictability (Henseler et al., 2009; Shmueli et al., 2016), but it is inappropriate for creating or improving theories. It is also suitable for smaller sample sizes, particularly non-normal distributions, which are ideal for formative and reflective complex formative particulars (Diamantopoulos & Riefler, 2011). The CB-SEM allows for the creation of a new theory, but it requires a larger sample size and normally distributed data. It is also appropriate when the reflective measures conflict with the theory during the formative model. The CB-SEM needs an adaptive model that can link the disturbance terms until the model with the data is adapted based on parameters such as RMSE χ^2. PLS-SEM seems to be the most convenient and fastest problem-solving method (Hair et al., 2012, p.415). However, Rönkkö and Everman (2013, p.19) argue that the PLS-SEM cannot define a theory, while Rönkkö et al. (2015, p.82) say that it should not be used in psychology studies as there is no error measurement (Rönkkö & Everman, 2013). Thus, PLS-SEM results are biased, even when the sample size increases to infinity (Kock, 2017), resulting in errors throughout the coefficient of the path, weight, and load. Dijkstra and Henseler (2015a,b) proposed consistent partial least squares (PLSc) to measure common factors or latent variables in order to solve the above problems. Thus, in the measurement model, PLS-SEM can be a measuring factor or a latent variable using PLSc. Many researchers find the problems of causal-formative variable-specific interpretation between the indicators and the model construction to be confounding (Bollen & Bauldry, 2011). Henseler et al. (2014) invented CCA and described it in Henseler (2017b), and Schuberth et al. (2018). Their construct variable consisted of a composite or emerging variable in the form of a combination of indicators. Thus, if all of the model's latent variables are CFA, then all of the model's emergent variables are CCA (Henseler, 2017a). This is a modernized PLS-SEM.

Mikko Rönkkö is a PhD student at Aalto University in Finland, and his dissertation focused on the PLS matrix. Rönkkö and Everman (2013) criticised the typical PLS-SEM technique, which utilizes composite data in both reflective latent and causal formative measurement models. Ten well-known PLS-SEM researchers felt compelled to respond to their critiques (Henseler et al., 2014). This is the most crucial information for the development of the modern PLS-SEM. Dijskstra and Henseler (2015ab) used the PLSc to quantify a latent reflective measurements model indicating that all constructs were PLSc, and hence, that it was a CFA. Henseler (2017a), Henseler (2017b), Henseler et al. (2014), Hubona et al. (2021), Muller et al. (2018), and Schuberth et al. (2018) suggested using CCA to discover a model's formative composite, which is based on the fact that all constructs are emergent variables, that is a pair of PLS-SEMs that transition from conventional to contemporary, as stipulated by Rönkkö and Everman (2013). The next part will discuss Henseler' s CCA.

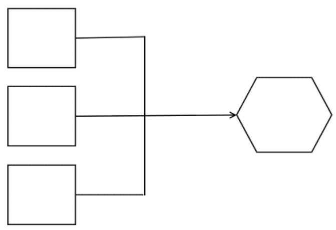
Confirmatory composite analysis

Concerning the CB-SEM, Jöreskog, (1969) suggested that the CFA has been more widely used in social sciences for the 50 prior years. Presently, the PLS-SEM can be utilized in both the CFA and the CCA. Several CFAs have already been investigated (e.g., Schumacker & Lomax, 2016), which is why this study focuses on CCAs. There are two conceptualization of CCA. The first is by Henseler et al. (2014) and Schuberth et al. (2018). Who are first developed CCA. The second by Hair et al. (2019), and Hair et al. (2020), uses the same name, but that is of no consequence. A CCA as defined by Hair et al. (2020) is when the measurement model is reliable and valid and can be used for all classical reflective, causal-formative, latent variables (PLSc) and composite variables. But there are at least three weaknesses in Hair et al. (2020) CCA: (1) it is not reliable for predictions, (2) it may use both reflective and formative measures, and (3) it does not have a goodness of fit index (Henseler & Schuberth, 2020). The researcher will use the original CCA for the study, which better benefits the academic society and preserves merit.

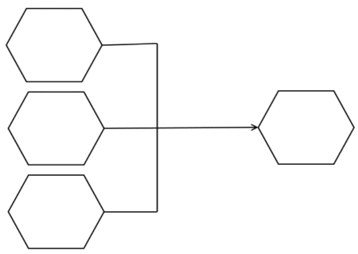
 Creative Business and Sustainability Journal (CBSJ)
Vol.44 No.2 July – December 2022, pp.22-40

Methods of constructing emergent variables

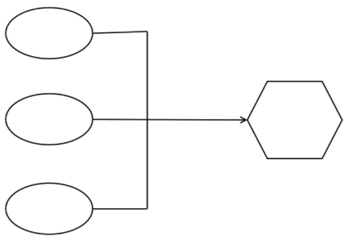
As discussed above, a CFA is created using constructs in the model that are exclusively latent variables, while a CCA is created using constructs that are all emergent variables. The latent variables are specific to an individual's attitudes, traits and behavior, while emergent variables have wider meanings covering things such as capabilities, values, indices, therapies, and interventions, as well as artifacts or design ideas (Hubona et al., 2021; Schubert et al., 2018). Additionally, emergent variables can be created from latent variables or attitudes, traits and behaviors. Yu et al. (2021) proposed a method to create emergent variables as shown in figure 1.



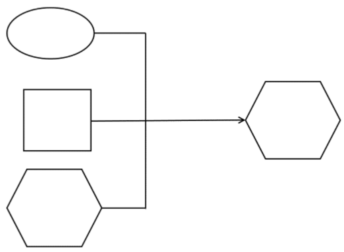
(a) An emergent variable made of observable variables (default)



(b) An emergent variable made of emergent variables



(c) An emergent variable made of latent variables



(d) An emergent variable made of different types of variables

Figure 1: Emergent Variables Made up of Different Components (Source: Yu et al., 2021)

Variables can be classified into three types: Observable variables, emergent variables, and latent variables. In figure 1, a square represents an observable variable, a hexagon represents an emergent variable, and an oval represents a latent variable. Naturally, observable and emergent variables can be constructed using the CCA method, whereas latent variables are constructed using the CFA method. CCA and CFA use a formative and reflective measurement model, respectively. However, this study needed to employ latent variables to accurately measure the formative variables depicted in Figure 1 (c). As illustrated in Figure 1 (d), the emergent variables were constructed by combining latent, observable, and emergent variables. The construct variables in Figure 1 (a), (b), (c), and (d) are all emergent variables, thus requiring a CCA.

Stages for constructing a CCA

The CCA is an SEM subtype (Schuberth, 2020) which uses the same four-stage CFA path of analysis: (1) composite model specification, (2) composite model identification, (3) composite model estimation, and (4) composite model assessment (Schuberth et al., 2018).

Specifications of the composite model

This stage is essential for the specific relationship between the composite variables, including their nomological-based indicators. The definition of a composite variable depends on the design of each research study, meaning the same composite variable differs from study to study (Sarstedt et al., 2016). Researchers, therefore, should prepare many such conceptual frameworks, composite variables, and their indicators before analysis. The CCA assumes, in general, that a composite variable is a linear combination of its indicators. The recursive combination of the composite variable and the endogenous and exogenous variables can be specified to summarize the error (Antonakis et al., 2010). Composite variables have certain drawbacks that must be measured in order to overcome multicollinearity between them using 2SLS estimates (Hult et al., 2018). Alternatively, each construct's indicators should not be closely linked, otherwise, a constructed variable of a higher-order should be created. For example, brand equity (Aaker, 1991) combines brand awareness, brand associations, brand quality, brand loyalty, and other proprietary assets. With the CCA, the indicator is not causal to the composite variables, but the indicator explains how to build the composite variables as an ingredient or composition (Henseler, 2017b). Latent variables have limitations on the relationship between the indicators and have internal consistency parameters that differ from composite variables (Henseler, 2017b, p181). In both the composite variables and the indicators, the composite variables have no error terms. There is no unit definition available for a composite variable, it refers to the description of a linear indicator combination (Bollen & Bauldry, 2011). However, some composite variables mean unity, especially in hierarchical constructs (Henseler, 2017b) with fewer indicators. Many development organizations, such as OECD, use a composite index to construct their indicators.

Identification of the composite model

Model identification links the model's specifications to the computer algorithm showing that the parameter has a unique solution (Raykov & Marcoulides, 2006). Model identification can provide consistent parameter estimation and reliable interpretation (Marcoulides & Chin, 2013). There are three types of model identification: under-identification, just-identification, and over-identification. The degree of freedom (df) is used to identify the type. Schuberth et al. (2018, p5) reported that an under-identification model offers multiple parameter sets that are consistent with the model's limitations, so there is no unique parameter solution for a degree of freedom below zero. It can only lead to dubious conclusions. A justified model provides a unique model parameter solution, and has as many free parameters as non-redundant matrix elements. When df is zero, the data fit is relatively perfect without assessment. An over-identified model has a unique solution but provides more non-redundant matrix elements than model parameters when df is more than zero. It can be used in empirical studies to evaluate the overall fit model as it should be within the sampling error limitations if valid.

Estimation of the composite model

To estimate a CCA, Schuberth et al. (2018) used the canonical correlation framework discussed in Dijkstra (2010) to evaluate the best relationship between two composite groups by generating canonical variables. The composite groups can have many indicators that are equal or not. Canonical correlation analysis transforms two multidimensional composite indicators into linear relationships by evaluating the weight of the matrix as being diagonal, and the diagonal correlation as being maximized.

Assessment of the composite model

Nomological validity[1], trust, and weight (composition) are the evaluation tools of the composite model (Henseler, 2017b). CCA provides the statistical technique needed to test the nomological validity of composites (Henseler et al. 2014). It is used when the saturated model implies such a small discrepancy between the empirical correlation matrix and the correlation matrix that the sampling error is attributable to this discrepancy (Henseler, 2017a). The nomological validity refers to the significance of the relationship of the composite variable as having a size, sign, and model fit. If the composite is nomologically valid, the researcher concludes that the composite is not an individual indicator, but acts within a nomological network (Henseler, 2017). The model fit measured uses three parameters: squared Euclidean distance (d_L), geodesic distance (d_G), and standardized root mean square residual (SRMR). The three parameters are fitted perfectly at zero (Schuberth et al. 2018). If the three parameters are less than the bootstrap-based percentiles of 95 percent ('HI95') or 99 percent ('HI99') to generate an empirical discrepancy distribution, then the model is valid. It is doubtful that the model will be correct if it exceeds these values (Henseler, 2017b). However, where this condition cannot be obtained, an SRMR below 0.08 will show a good fit (Hu & Bentler, 1998). That factor variable criteria, however, needs to be investigated for the composite variables (Henseler, 2017b).

As for reliability, it is possible to measure the composite variable using observed variables based on nomological validity. No random measurement error occurs, and the resulting composite reliability is 1. Also, if there is an unexpected measuring error in the indicator, the composite variable isnot compleately reliable (Henseler, 2017b). As for weightings, Henseler (2017b) stated that one must carefully consider if the analysis does not have significance but freely figures weights, meaning the size, sign and confidence intervals must be determided. Additionally, Henseler (2017a) argued that in order to account for multicollinearity in CCA, one should assess the VIF and guarantee that it is less than five, and ideally, less than 3.3 for detecting common method bias (Kock, 2015).

Data, conceptual framework and hypothesis

Latent variables such as satisfaction, trust, and loyalty have been suggested as emergent variables for the CCA. The CFA was created to facilitate comparison with the CCA findings.

[1] Cronbach and Meehl (1955) established a valid nomological network. It defines the construct validity, which consists of three concepts: Observable aspects, theoretical constructs, and relationships among and between observable items and theoretical constructs. But it is mainly used for construct validity. It is used for model specification when the two constructs are associated with other constructs in the nomological network, indicating that the constructs are valid, including those linked to other indicators and constructs.

The following hypothesis and conceptual framework was constructed, in which, diagrammartically, ovals are considered latent variables and hexagons are considered emergent variables. According to numerous studies, satisfaction has a positive effect on trust (Geyskens et al., 1999), while trust has a positive effect on loyalty (Chaudhuri & Holbrook, 2001) and satisfaction has a positive effect on loyalty (Ehigie, 2006). Recent research has demonstrated that customer satisfaction does not always have a direct effect on customer loyalty, but frequently works through mediators (Izogo, 2015; Picon et al., 2014). This study uses trust as a mediator variable between satisfaction and loyalty (Hart & Johnson, 1999). In this study, trust will function as a mediator if a and b are significant, with a full mediation effect when c is insignificant and a partial mediation effect when c is significant (Nitzl, et al., 2016) as follows:

H1: Satisfaction had a positive effect on trust.
H2: Trust had a positive effect on loyalty.
H3: Satisfaction had a positive effect on loyalty.
H4: There are mediation effects between satisfaction and loyalty.

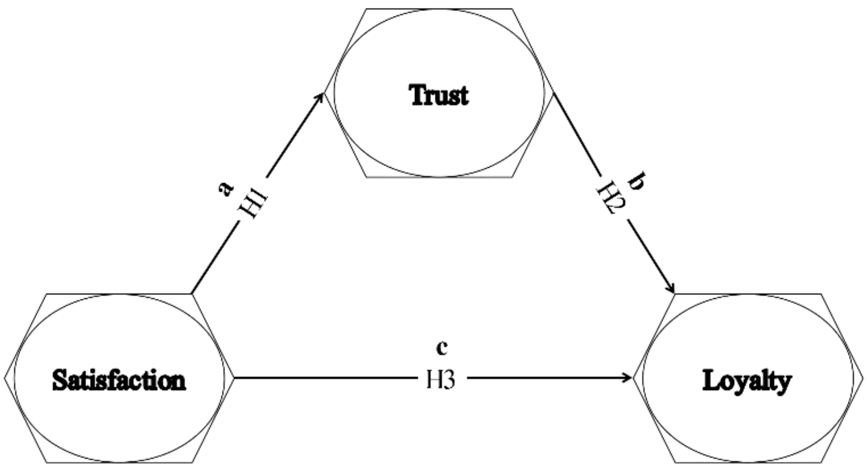


Figure 2: The Conceptual Framework[1]

Research Methodology

Sample and Population

The sample size was established using Soper (2021) approach, which considers the effect size (0.15), power (0.80), latent variables (3), indicators (9), and probability level (0.05). As a result, a minimum sample size of 200 students was selected out of a population of 1,464. Using a Google form, the data were collected randomly from students of Valaya Alongkorn Rajabhat University. As a result of COVID-19, the majority of the students were studying online, facilitating the process.

[1] This diagram illustrates two ways in which all ovales are latent variables that are analyzed using a CFA. All hexagons are emergent or composite models requiring a CCA.

Questionnaire

The questionnaire used was divided into two sections, the first of which comprised the sample's biographical information: Gender, year of study, and income. Women accounted for almost 75% of the population. The students were enrolled in either their 1st, 2nd, 3rd, or 4th year, accounting for around 20.50%, 51.50%, 17%, and 11%, respectively. Three monthly income levels were set: less than 4,000, 4001-7000, and more than 7000 baht. The results reveal the distribution to be about 49.50%, 36%, and 14.50%, respectively. The second part of the questionnaire addressed the indicator, which this study defined as follows: From Zeng et al. (2009), the indicator of satisfaction is adaptive, from Parasuraman et al., (2005), loyalty is adaptable and from Urban et al. (2009), the trust indicator is inspired. The customer satisfaction questionnaire inquired about product usability, product suitability, and promotion. The customer loyalty survey inquired about future purchases, recommendations to friends, and brand ratings.The trust questions were related to tracking of purchase progression, usage security, and customer data collection security. The reliability figures for satisfaction, loyalty, and trust questions were 0.81, 0.86, and 0.87, respectively.

Table1 Source of Indicators

Indicators		Items	Sources
Satisfaction			
	Satis1	The online shopping applications for Lazada and Shopee are simple to use.	Adapted from (Zeng et al. (2009)
	Satis2	The applications for Shopee and Lazada allow me to purchase high-quality products that fit my needs.	
	Satis3	Lazada and Shopee applications are often used to promote and sell items, typically with discounts and free shipping.	
Loyalty			Adapted from, Parasuraman et al (2005)
	Loyal1	My next purchase will be made using the Lazada or Shopee application.	
	Loyal2	I will recommend the Lazada or Shopee application to my relatives and friends.	
	Loyal3	For the Lazada or Shopee application, I will offer the maximum possible rating.	
Trust			Inspired by Shankar et al (2002) and Urban et al (2009)
	Trust1	I can track the progression of my purchases when I utilize the Lazada or Shopee application.	
	Trust2	Lazada and Shopee's application are safe to use and orders can be verified.	
	Trust3	The client information collected by the Lazada and Shopee applications is secure and safe.	

The quality of CFA & CCA

The quality of a PLS-SEM is based on the parameters of model fit, measurements and structural models. However, there are differences in the quality of a measurement model, therefore, a saturation and an estimation model were used to assess the model's fit quality, while a bootstrapping model was used to identify discrepancies between the data and the model's suggested correlation matrix Henseler (2017a). Three statistics were used to evaluate model fit: The standardized root mean square residual (SRMR), the unweighted minimum square discrepancy (d_{ULS}), and the geodesic discrepancy (d_G). The smaller the outcomes, the better. Two prerequisites apply: To begin, the outcomes of the quantile bootstrapping at the 95th and 99th percentiles (HI95 and HI99) should be less than the requirements. Second, if the first criterion is not met, the SRMR value, which should be less than 0.08, may need to be examined (Hu & Bentler, 1998).

The measurement model consisted of the latent reflective models and the composite formative models. The parameters of the latent reflective models consisted of internal consistency reliability, indicator reliability, convergent validity, and discriminant validity. Internal consistency is defined by two parameters: Dijkstra-Henseler's (A) rho and the Jöreskog's (C) rho, which should both be greater than 0.70. Indicator loading should be more than 0.708 (Disjkatra & Henseler, 2015a), suggesting that it may be used to infer the structure of the item. Convergent validity indicates that the indicator is strongly related to extracted average variance (AVE) larger than 0.50 within the same framework. The discriminant validity criterion is based on the presence of HTMT2, which must be distinct and less than 0.85 (Henseler, 2017a).

For the measurement of the composite model, nomological validity, multicollinearity, weight relevance, and reliability were all taken into account when determining the quality of the composite formative model. Nomological validity is a theory-based assessment of an indicator's sign and magnitude that is heavily weighted during the model fit stage. This is largely dependent on the specification model, which may include a multicollinear issue. The researcher can omit some indicators with very large variation inflation factors (VIFs). This may help to increase nomological validity, as well as to solve multicollinearity and overall fit problems. Thus, this is critical in practice for the researcher. The multicollinearity test requires that the variance inflation factors (VIFs) not exceed 3.3 for detecting common method bias (Kock, 2015). For weighting, if the indicator weightings are not significant, their loading must be above 0.50 for the indicator to remain within the model.

Both latent reflective and emergent formative structural models may use similar quality criteria. There are three components: The effect size (f^2), the relationship size (R^2), and the predictability size (Q^2). Effect sizes of 0.02, 0.15, and 0.35 may be categorized as small, medium, and substantial, respectively. Small, medium, and substantial relationships, sizes are categorized as 0.25, 0.50, and 0.75 respectively. All the results may be considered statistical significance if the t-statistic value is equal to or more than 1.96.

Research Findings

Total model fit

Table 2 demonstrates that the three model fit parameters, SRMR, d_{ULS}, and d_G, are all smaller than the 95 percent bootstrap-based percentile ('HI95') for both the CFA and the CCA. Due to the fact that the reference data and model are similar, the model fit parameters for the CCA are somewhat better than those of the CFA. Additionally, the CCA model fit results reveal that the indicators and constructs have nomological validity, meaning that their sizes and signs are compatible with the theory.

Table 2: Total Model Fit

Parameters		Saturated Model			Estimated Model		
		Value	HI95	HI99	Value	HI95	HI99
CFA	SRMR	0.028	0.034	0.037	0.028	0.034	0.037
	d_{ULS}	0.036	0.051	0.062	0.036	0.051	0.062
	d_G	0.042	0.070	0.087	0.042	0.070	0.087
CCA	SRMR	**0.027**	**0.029**	**0.033**	**0.027**	**0.029**	**0.033**
	d_{ULS}	0.033	0.038	0.049	0.033	0.038	0.049
	d_G	0.039	0.054	0.068	0.039	0.054	0.068

Measurement model

The CFA model's measurement model parameters are shown in Table 3. The internal consistency analysis demonstrates that all latent variables have values larger than 0.7 for **rho (ρ_A)**, rho (ρ_c), and alpha, suggesting construct validity. All indicators have a loading larger than 0.708, indicating that the loading squares or reliability indicators of satisfaction, trust, and loyalty are all above the 0.50 threshold. All convergent validity coefficients are above 0.50. However, Table 4 shows that the CFA analysis of this model indicates a problem with the discriminant validity when HTMT2 is used, since all latent variables in the table have a value larger than 0.85. The heterotrait-monotrait ratio (HTMT) is based on tau-equivalent measurement models, which are unlikely to hold true in most empirical studies. In comparison to the HTMT, the HTMT2 yields less biased estimations of the correlations among the latent variables, especially if the loading patterns of the indicators are varied. So, when there are congeneric measurement models, the HTMT2 should be used instead of the HTMT to check for discriminant validity (Roemer, Schuberth, & Henseler, 2021)

Table 3: Measurement Model Parameters

Indicator	Loading	Reliability Indicator	Dijkstra-Henseler's rho (ρ_A)	Jöreskog's rho (ρ_c)	Cronbach's alpha (α)	Convergent Validity (AVE)
Satisfaction			0.813	0.812	0.812	0.590
Satis1	0.728	0.529				
Satis2	0.798	0.637				
Satis3	0.776	0.603				
Loyalty			0.865	0.865	0.865	0.681
Loyal1	0.806	0.650				
Loyal2	0.848	0.718				
Loyal3	0.820	0.673				
Trust			0.867	0.866	0.866	0.683
Trust1	0.807	0.650				
Trust2	0.836	0.699				
Trust3	0.837	0.700				

Creative Business and Sustainability Journal (CBSJ)
Vol.44 No.2 July – December 2022, pp.22-40

Table 4: Discriminant Validity (HTMT2)

Construct	Satisfaction	Loyalty	Trust
Satisfaction			
Loyalty	0.923		
Trust	0.917	0.901	

Table 5 shows the quality parameters of the CCA which are the nomological validity, significance of weighting, multicollinearity and reliability. The overall model fit index indicates that the model is nomologically valid, while reliability is fixed at one. All indicators have a t-statistic value greater than 1.96, indicating they are all statistically significant. The VIFs range from 1.724 to 2.281, which is significantly less than three, indicating the absence of problems with multicollinearity and common method bias.

Table 5 Measurement Model of CCA

Indicator	Mode B			
	Loading	**Weight**	**t-statistic of Weight**	**VIF**
Satis1	0.801	0.280	3.010	1.777
Satis2	0.885	0.489	5.697	1.724
Satis3	0.859	0.399	4.599	1.844
Loyal1	0.864	0.295	3.877	2.279
Loyal2	0.904	0.427	4.045	2.281
Loyal3	0.889	0.403	5.253	2.138
Trust1	0.863	0.303	3.487	2.224
Trust2	0.899	0.414	4.970	2.234
Trust3	0.899	0.407	5.405	2.281

Comparing the measurement models of the CFA and the CCA, the former's discriminant validity of satisfaction, trust, and loyalty as latent variables is highly questionable. Thus, the CCA measuring paradigm is more successful than that of the CFA for all items.

Structural model

Table 5 illustrates the structural model quality of the CFA using the path coefficient, the t-statistic, the f^2 statistic, and the R^2 statistic. In the reject hypothesis, the two paths have no significance. There is significance on the path from satisfaction to trust, which leads to the acceptance of the hypotheses. The t-statistics for the three paths are above 1.96 in the CCA model, indicating significance. All three paths thus support the hypotheses. The results demonstrate that all parameters, including f^2, R^2, path coefficient, and even Q^2, are less significant than the t-statistic in this context. Figures 3 and 4 illustrate the data given in Table 6.

Table 6 Hypothesis of CFA and CCA

	The Effect	Beta	Standard error	t-value	p-value	Cohen's f^2	Hypothesis
CFA	H1: Satisfaction -> Trust	0.918	0.041	22.348	0.000	5.321	Supported
	H2: Trust -> Loyalty	0.310	4.894	0.063	0.949	0.126	Not supported
	H3: Satisfaction -> Loyalty	0.645	4.893	0.132	0.895	0.546	Not supported
CCA	H1: Satisfaction -> Trust	0.768	0.036	21.628	0.000	1.437	Supported
	H2: Trust -> Loyalty	0.436	0.073	5.931	0.000	0.258	Supported
	H3: Satisfaction -> Loyalty	0.453	0.070	26.616	0.000	0.278	Supported

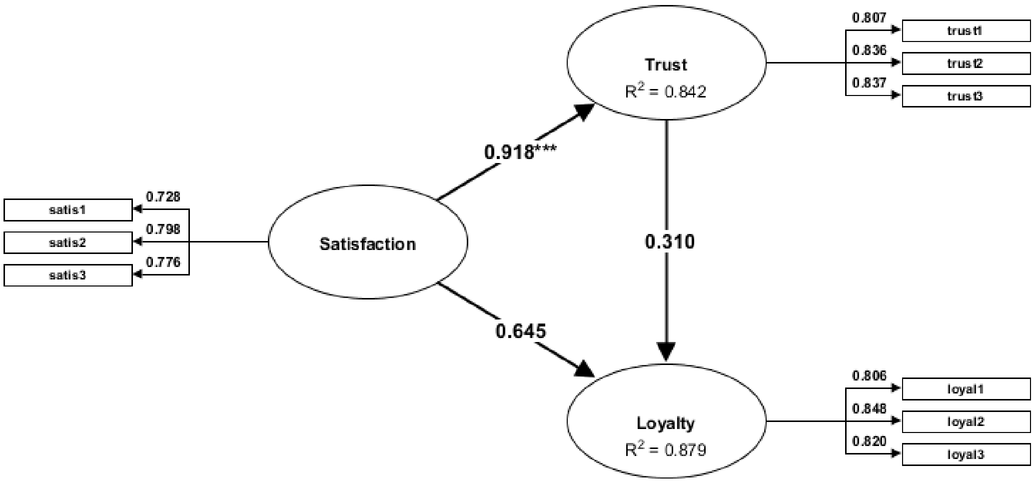


Figure 3: CFA Results

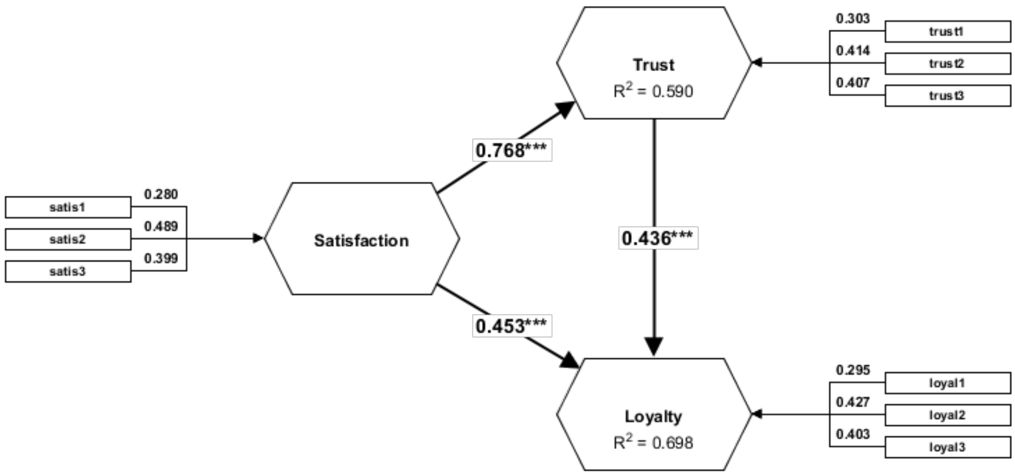


Figure 4: CCA Results

The CCA is more effective than the CFA in the structural model, as its results are significant along all three paths. On the other hand, the CFA results are significant along only one path.

Mediation effect

Table 7 shows that the CFA has a negligible mediation impact. Additionally, there are indirect relationships between satisfaction and loyalty in the CCA. The indirect impact is 0.768 x 0.436, or 0.335, and there is a significant link between satisfaction and loyalty. As a result, there is a partial mediation effect, with trust acting as a mediator.To calculate the t statistic, the coefficient is divided by its standard error.

Table 7 Mediation Effect

		Indirect Effect	Standard Error	t-value	p-value	Hypothesis
CFA	H4: Satisfaction -> Loyalty	0.285	4.892	0.058	0.954	Not supported
CCA	H4: Satisfaction -> Loyalty	0.335	0.059	5.701	0.000	Supported

Discussion

While latent variables may be drawn from attitudes, traits, and behaviors, emergent variables can be derived from capabilities, indices, values, therapies, and interventions (Hubona et al. 2021; Schuberth et al, 2018). This study demonstrates that emergent variables may be created from latent variables, as stated by Hubona et al. (2021).

The common factor is used for behavior analysis, while the composite variable is composed of several types of variables depending on the design. A CB-SEM can only be utilized in CFA while a PLS-SEM can be used in both CFA and CCA, providing researchers with benefits such as fast convergence and the use of smaller sample sizes. However, small sample sizes can be less accurate than larger ones (Rigdon, 2016), and as such, their use should be carefully considered (Benitez et al., 2020). This study reconfirms that the PLS-SEM is advantageous when using a small sample size, while at the same time offering the required statistical power that CB-SEM provides. In addition, the total model fit index, the structural model and indirect effect obtained in this study indicate that CCA is more effective than the CFA.

Managerial Implication

CCA proved to be more adaptable and precise than CFA, with all paths between the emergent variable of the CCA and the indirect effects being significant. The CFA in this study may have been weakened by the small sample size, since Soper (2021) suggested sample sizes of 200 for model structure and 545 for detecting effects. Thus, the studies need utilize sufficiently large sample size for a CFA when studying indirect effects. As a result, it may be preferable to use a CCA to study the mediation effects when using a smaller sample size. Subsequent CCA studies are needed, and should emergent variables be missing, then the researcher can create them from latent variables.

The findings of this study also suggest that trust mediates the impact between consumer satisfaction and customer loyalty in online shopping platforms.

Theoretical Implications

This study implies that emergent variables can be constructed from latent variables, which is remarkable given the scarcity of literature on this topic. Jhantasana (2022) is the only researcher who had previously constructed emergent variables from latent variables relating to extrinsic and intrinsic motivation, and job satisfaction. Our results are in line with those of Jhantasana (2022), who found that building emergent variables from latent variables for CCA is feasible and of higher quality than CFA. As a result, theoretical issues about creating emergent variables from latent variables have been proven to be possible, as proposed by Hubona et al (2021) and implemented by Jhantasana (2022). This study also confirms the theory that trust acts as a powerful mediator between customer satisfaction and customer loyalty when it comes to online shopping platforms.

Conclusion

CFA using CB-SEMs has been used in behavioral research for over 50 years with resulting gaps in the emergent variables. PLS-SEMs, on the other hand, may be used for many types of research, including confirmatory, exploratory, and predictive (Henseler 2018; Henseler et al., 2014; Shmueli et al., 2016;). Thus, the PLS-SEM is a more effective statistical tool for doing empirical research in a variety of business and social science disciplines (Benitez et al., 2020). PLS-SEMs can be used to perform CFA and CCA, utilizing latent and emergent variables, respectively. According to Hubona et al. (2021), this form of research, which creates emergent variables from latent ones for CCA, yields higher-quality results than CFA. Our study confirmed this using Lazada and Shopee online platforms, with trust serving as a partial mediator between customer satisfaction and consumer loyalty.

Limitations and Future Research

While this study provided tremendous advancements, some limitations should be stated. The sample size was relatively small, which caused the CFA to introduce bias. Additionally, even though students are generally responsible consumers who respond to new technology, the research would have benefited from the use of a sample that included people from all walks of life. The results would likely have produced a more accurate representation of general consumer behavior. Additionally, future studies should be designed to include other areas or academic specialties in order to construct emergent variables from latent variables.

Acknowledgements: We would like to convey our gratitude to those reviewers who helped ensure the accuracy and efficacy of our work. Any errors or omissions are solely the responsibility of author. We also would like to acknowledge the help provided by Bachelor of Accountancy students Tippawan Taweekasikam, Panithan Kamnak, and Piyawan Kherassa who assisted with the data collection.

References

Aaker, D. A., & Equity, M. B. (1991). *The free press*. New York.

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, *21*(6), 1086-1120.

Arnett, D. B., Laverie, D. A., & Meiers, A. (2003). Developing parsimonious retailer equity indexes using partial least squares analysis: A method and applications. *Journal of Retailing*, *79*(3), 161-170.

Benitez, J., Henseler, J., Castillo, A., & Schuberth, F. (2020). How to perform and report an impactful analysis using partial least squares: Guidelines for confirmatory and explanatory IS research. *Information & Management*, *57*(2), 103-168.

Bollen, K. A., & Bauldry, S. (2011). Three Cs in measurement models: Causal indicators, composite indicators, and covariates. *Psychological Methods*, *16*(3), 265.

Cadogan, J. W., & Lee, N. (2013). Improper use of endogenous formative variables. *Journal of Business Research*, *66*(2), 233-241.

Chaudhuri, A., & Holbrook, M. B. (2001). The chain of effects from brand trust and brand affect to brand performance: The role of brand loyalty. *Journal of Marketing, 65*, (2), 81-93.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281.

Diamantopoulos, A., & Riefler, P. (2011). Using formative measures in international marketing models: A cautionary tale using consumer animosity as an example. *Advance in International Marketing,22, 11-30*.

Dijkstra, T. K., & Henseler, J. (2015a). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics & Data Analysis*, *81*, 10-23.

Dijkstra, T. K., & Henseler, J. (2015b). Consistent partial least squares path modeling. *MIS Quarterly*, *39*(2), 1-A5

Ehigie, B. O. (2006). Correlates of customer loyalty to their bank: a case study in Nigeria. *International Journal of Bank Marketing, 24* (7), 494-508.

Geyskens, I., Steenkap, J. E. B. M., & Kumar, N. (1999). A Meta-analysis of satisfaction in marketing channel. *Journal of Marketing Research, 36*(2), 223-238

Guenzi, P., Johnson, M. D., & Castaldo, S. (2009). A comprehensive model of customer trust in two retail stores. *Journal of Service Management*, *20*(3), 290-316.

Hart, C. W. & Johnson, M. D. (1999). Growing the trust relationship. *Marketing Management, 8* (1), 8-19.

Hair Jr, J. F., Howard, M. C., & Nitzl, C. (2020). Assessing measurement model quality in PLS-SEM using confirmatory composite analysis. *Journal of Business Research*, *109*, 101-110.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis: International version*. New Jersey, Pearson.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis, cengage learning*, *andover*. Hampshire, United Kingdom.

Henseler, J. (2017a). *ADANCO 2.0.1.*: *user manual*. https://www. composite-modeling.com/ support/user-manual/

Henseler, J. (2017b). Bridging design and behavioral research with variance-based structural equation modeling. *Journal of Advertising*, *46*(1), 178-192.

Henseler, J., & Schuberth, F. (2020). Using confirmatory composite analysis to assess emergent variables in business research. *Journal of Business Research*, *120*, 147-156.

Henseler, J., Ringle, C., & Sinkovics, R. (2009). The use of partial least squares path modeling in international marketing. *Advance in International Marketing, 20*, 277-319.

Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., Straub, D. W., Ketchen, D., Hair, J. F., Hult, T. M., & Calantone, R. J. (2014). Common beliefs and reality about PLS: Comments on Ronkko and Evermann (2013). *Organizational Research Methods, 17*(2), 182-209.

Horppu, M., Kuivalainen, O., Tarkiainen, A. & Ellonen, H.-K. (2008). Online satisfaction, trust and loyalty and the impact of the offline parent brand. *Journal of Product & Brand Management*, *17*(6), 403-413.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to under parameterized model misspecification. *Psychological Methods*, *3*(4), 424.

Hubona, G. S., Schuberth, F., & Henseler, J. (2021). A clarification of confirmatory composite analysis (CCA). *International Journal of Information Management*, *61*, 102399.

Hult, G. T. M., Hair Jr, J. F., Proksch, D., Sarstedt, M., Pinkwart, A., & Ringle, C. M. (2018). Addressing endogeneity in international marketing applications of partial least squares structural equation modeling. *Journal of International Marketing*, *26*(3), 1-21.

Izogo, E. E. (2015). Determinants of attitudinal loyalty in Nigerian telecom service sector: Does commitment play a mediating role?. *Journal of Retailing and Consumer Services, 23*, 107-117.

Jhantasana, C. (2022). Intrinsic and extrinsic motivation for university staff satisfaction: Confirmatory composite analysis and confirmatory factor analysis. *Asia Social Issues*, *15*(2), 249810.

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika, 34*, 183-202.

Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of e-Collaboration*, *11*(4), 1-10.

Kock, N. (2017). *Factor-based SEM building on consistent PLS: An information systems illustration.* Laredo, TX: ScriptWarp Systems.

Marcoulides, G. A., & Chin, W. W. (2013). You write, but others read Common methodological misunderstandings in PLS and related methods. In *Ed New Perspectives in partial least squares and related methods* (pp. 31-64). Springer, New York, NY.

Müller, T., Schuberth, F., & Henseler, J. (2018). PLS path modeling–a confirmatory approach to study tourism technology and tourist behavior. *Journal of Hospitality and Tourism Technology*. 1757-9880.

Nitzl, C. (2016). The use of partial least squares structural equation modeling (PLS-SEM) in management accounting research: Directions for future theory development. *Journal of Accounting Literature*, *37*, 19-35.

Nitzl, C., Roldan, J. L., & Cepeda, G. (2016). Mediation analysis in partial least squares path modeling: Helping researchers discuss more sophisticated models. *Industrial Management & Data Systems, 116*(9). 1849-1864.

OECD. (2006). Handbook on Constructing Composite Indicators. https://www.oecd.org/sdd/42495745.pdf

Parasuraman, A., Zeithaml, V. A., & Malhotra, A. (2005). ES-QUAL: A multiple-item scale for assessing electronic service quality. *Journal of Service Research*, *7*(3), 213-233.

Picon, A., Castro, I. & Roldan, J. L. (2014). The relationship between satisfaction and loyalty: A mediator analysis. *Journal of Business Research, 67*(5), 746-751.

Raykov, T. & Marcoulides, G. A. (2006). *A First Course in Structural Equation Modeling*, 2nd Edn. Mahaw: Lawrence Erlbaum Associates.

Reinartz, W., Krafft, M., & Hoyer, W. D. (2004). The customer relationship management process: Its measurement and impact on performance. *Journal of Marketing Research*, *41*(3), 293-305.

Richter, N. F., Cepeda-Carrión, G., Roldán Salgueiro, J. L., & Ringle, C. M. (2016). European management research using partial least squares structural equation modeling (PLS-SEM). *European Management Journal, 34* (6), *589-597.*

Rigdon, E. E. (2016). Choosing PLS path modeling as an analytical method in European management research: A realist perspective. *European Management Journal*, *34*(6), 598-605.

Rigdon, E. E., Sarstedt, M., & Ringle, C. M. (2017). On comparing results from CB-SEM and PLS-SEM: Five perspectives and five recommendations. *Marketing: ZFP – Journal of Research and Management, 39*(3), 4-16.

Ringle, C. M., Sarstedt, M., Mitchell, R., & Gudergan, S. P. (2018). Partial least squares structural equation modeling in human resource management research, *The International Journal of Human Resource Management, 31*(12),1617-1643.

Roemer, E., Schuberth, F., & Henseler, J. (2021). HTMT2–an improved criterion for assessing discriminant validity in structural equation modeling. *Industrial Management & Data Systems, 121*(12)*,* 2637-2650.

Rönkkö, M., & Evermann, J. (2013). A critical examination of common beliefs about partial least squares path modeling. *Organizational Research Methods, 16*(3), 425-448.

Rönkkö, M., McIntosh, C. N., & Antonakis, J. (2015). On the adoption of partial least squares in psychological research: Caveat emptor. *Personality and Individual Differences*, *87*, 76-84.

Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., & Gudergan, S. P. (2016). Estimation Issues with PLS and CBSEM: Where the bias lies!. *Journal of Business Research*, *69*(10), 3998-4010.

Sarstedt, M., Hair Jr, J. F., Cheah, J. H., Becker, J. M., & Ringle, C. M. (2019). How to specify, estimate, and validate higher-order constructs in PLS-SEM. *Australasian Marketing Journal (AMJ)*, *27*(3), 197-211.

Schuberth, F. (2020). Confirmatory composite analysis using partial least squares: Setting the record straight. *Review of Managerial Science*, 1-35.

Schuberth, F., Henseler, J., & Dijkstra, T. K. (2018). Confirmatory composite analysis. *Frontiers in Psychology*, *9*, 2541.

Schumacker., R. E., & Lomax, R. G. (2016). *A beginner's guide to structural equation* modeling*: 4th Edition*. New York, NY: Routledge.

Shankar, V., Urban, G. L., & Sultan, F. (2002). Online trust: A stakeholder perspective, concepts, implications, and future directions. *The Journal of Strategic Information Systems*, *11*(3-4), 325-344.

Shmueli, G., Ray, S., Velasquez Estrada, J. M. & Chatla, S. B. (2016). The elephant in the room: Evaluating the predictive performance of PLS models. *Journal of Business Research, 69*(10), 4552-4564.

Soper, D. S. (2021). *A-priori sample size calculator for structural equation models* [Software]. https://www.danielsoper.com/statcalc

Tran, T. T. (2019). On the factors affecting the development of e-commerce in Vietnam: Case study of Lazada, Shopee, and Tiki. *International Journal of Advanced and Applied Sciences, 6*(4), 45-52.

Urban, G. L., Amyx, C., & Lorenzon, A. (2009). Online trust: State of the art, new frontiers, and research potential. *Journal of Interactive Marketing*, *23*(2), 179-190.

Yoon, S.-H. (2007). Determinants of online service satisfaction and their impacts on behavioral intentions. *Journal of Korea Trade*, 11(3), 23-52.

Yoon, S.-J. (2002). The antecedents and consequences of trust in online-purchase decisions. *Journal of Interactive Marketing*, 16(2), 47-63.

Yu, X., Zaza, S., Schuberth, F., & Henseler, J. (2021). Counterpoint: Representing forged concepts as emergent variables using composite-based structural equation modeling. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, *52*(SI), 114-130.

Zeng, F., Hu, Z., Chen, R., & Yang, Z. (2009). Determinants of online service satisfaction and their impacts on behavioural intentions. *Total Quality Management & Business Excellence,* 20(9), 953-969.