UNIVERSITY OF TWENTE.

Webinar Faizan Ali

Confirmatory Composite Analysis

What is it & how is it applied?

Florian Schuberth[1]

[1]University of Twente

October 8, 2020

Overview

1. History of Confirmatory Composite Analysis
2. Confirmatory Composite Analysis
   - Model Specification
   - Model Identification
   - Model Estimation
   - Model Assessment
3. Demonstration

Who am I?

Florian Schuberth
email: f.schuberth@utwente.nl
web: www.florianschuberth.com

- 2007 - 2010: Bachelor in Business Administration & Economics; University of Würzburg, GER
- 2010 - 2012: Master in Economics; University of Würzburg, GER
- 2012 - 2017: PhD in Econometrics (M. Kukuk & J. Henseler); University of Würzburg, GER
- 2017 - present: Assistant professor; Product-Market Relations at the University of Twente, NL

History of CCA

In 2014, CCA was first mentioned as an approach to assess the overall fit of a composite model[1]

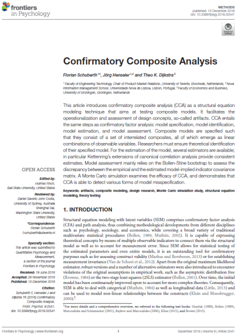
> well does PLS perform if the composite factor model is indeed correct?" Preliminary findings give rise to optimism for both these questions: PLS serves well for predictive purposes (Becker, Rai, & Rigdon, 2013), and when the composite factor model is true, PLS clearly outperforms covariance-based SEM and regression on sum scores (Rai et al., 2013). In conclusion, PLS is not a panacea, but certainly an important technique deserving a prominent place in any empirical researcher's statistical toolbox.
>
> **Authors' Note**
>
> The formal development of the composite factor model is attributable to the second author, and in collaboration with the first author developed into the concept of confirmatory composite analysis. The authors thank Dr. Rainer Schlittgen, Professor of Statistics and Econometrics, University of Hamburg, Germany, for helpful comments on earlier versions of the article.
>
> **Declaration of Conflicting Interests**
>
> The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

---

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

History of CCA

In 2018, CCA was fully elaborated[2] and in 2020, it was introduced to business research[3]



frontiers in Psychology

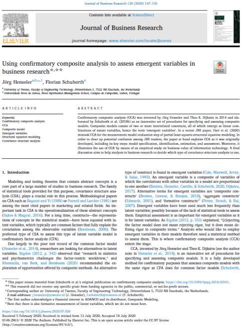
METHODS

Confirmatory Composite Analysis

Florian Schuberth, Jörg Henseler and Theo K. Dijkstra

OPEN ACCESS

1. INTRODUCTION



Journal of Business Research 120 (2020) 147–156

Journal of Business Research

Using confirmatory composite analysis to assess emergent variables in business research

Jörg Henseler, Florian Schuberth

ARTICLE INFO

ABSTRACT

1. Introduction

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

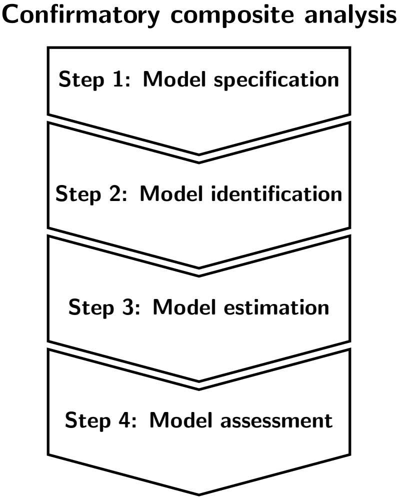
[3] Henseler, J., & Schuberth, F. (2020). Using confirmatory composite analysis to assess emergent variables in business research. *Journal of Business Research*, *120*, 147–156.

In memory and special thanks

Theo K. Dijkstra
*17.06.1953 - †14.09.2020

Confirmatory composite analysis

Like all approaches to SEM, CCA consists of four steps



Confirmatory composite analysis

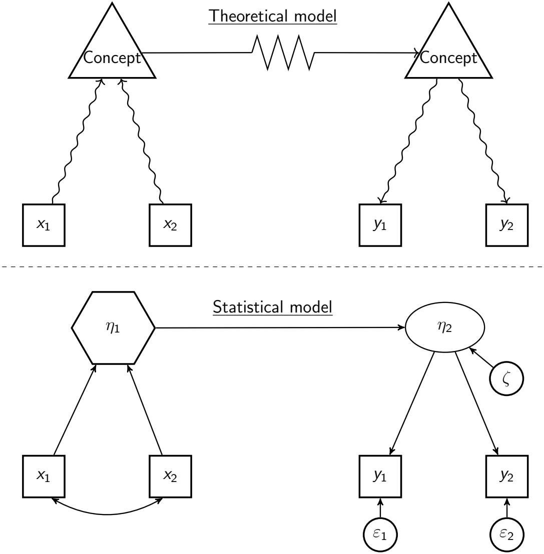
Step 1: Model specification

Step 2: Model identification

Step 3: Model estimation

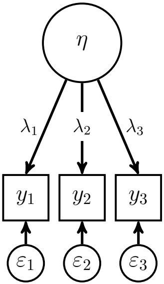
Step 4: Model assessment

Model specification



Two types of concepts

Type of concept:	Theoretical concept
Type of construct:	Latent variable
Dominant statistical model:	Common factor model

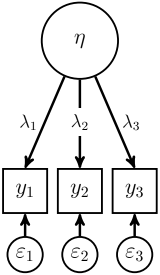


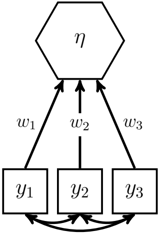
Statistical approach: Confirmatory factor analysis

Examples: Attitudes, traits

Two types of concepts

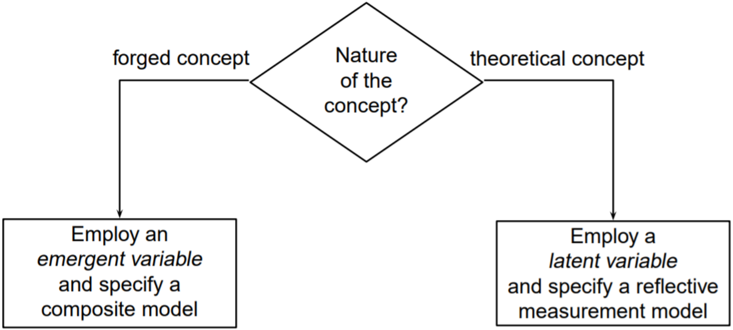
Type of concept:	Theoretical concept	Forged concept
Type of construct:	Latent variable	Emergent variable
Dominant statistical model:	Common factor model	Composite model





Statistical approach:	Confirmatory factor analysis	**Confirmatory composite analysis**
Examples:	Attitudes, traits	Capabilities, indices, values

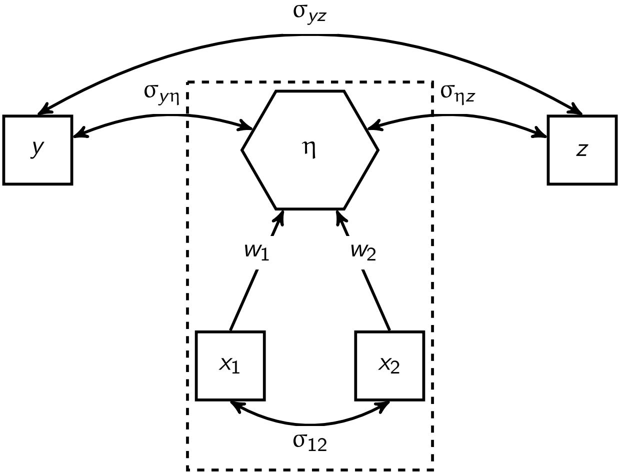
Decision



Decision about the outer model[4]

[4]Henseler, J. (in press). *Composite-based structural equation modeling: Analyzing latent and emergent variables*. New York, The Guilford Press.

Example specification



Minimal composite model

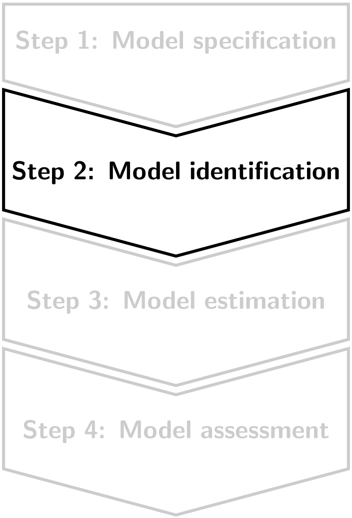
Model-implied variance-covariance matrix

The model-implied variance-covariance matrix $\boldsymbol{\Sigma}(\theta)$ of the minimal composite model:

$$\boldsymbol{\Sigma}(\theta) = \begin{array}{c} \begin{array}{cccc} \underline{y} & \underline{x_1} & \underline{x_2} & \underline{z} \end{array} \\ \begin{pmatrix} \sigma_{yy} & & & \\ \lambda_1\sigma_{y\eta} & \sigma_{11} & & \\ \lambda_2\sigma_{y\eta} & \sigma_{12} & \sigma_{22} & \\ \sigma_{yz} & \lambda_1\sigma_{\eta z} & \lambda_2\sigma_{\eta z} & \sigma_{zz} \end{pmatrix} \end{array} ,$$

where $\lambda_1 = \text{cov}(x_1, \eta)$ and $\lambda_2 = \text{cov}(x_2, \eta)$.

Confirmatory composite analysis



Model identification

Parameters in $\boldsymbol{\Sigma}(\theta)$ can be uniquely retrieved from the population variance-covariance matrix $\boldsymbol{\Sigma}$.

For the minimal composite model:

$$
\begin{array}{cccc}
\underline{y} & \underline{x_1} & \underline{x_2} & \underline{z} \\
\end{array}
\qquad\qquad
\begin{array}{cccc}
\underline{y} & \underline{x_1} & \underline{x_2} & \underline{z} \\
\end{array}
$$

$$
\begin{pmatrix}
\sigma_{yy} & & & \\
\lambda_1\sigma_{yc} & \sigma_{11} & & \\
\lambda_2\sigma_{yc} & \sigma_{12} & \sigma_{22} & \\
\sigma_{yz} & \lambda_1\sigma_{cz} & \lambda_2\sigma_{cz} & \sigma_{zz}
\end{pmatrix}
=
\begin{pmatrix}
\sigma_{yy} & & & \\
\sigma_{y1} & \sigma_{11} & & \\
\sigma_{y2} & \sigma_{12} & \sigma_{22} & \\
\sigma_{yz} & \sigma_{1z} & \sigma_{2z} & \sigma_{zz}
\end{pmatrix}
$$

(1) $\sigma_{yy} = \sigma_{yy}$ (2) $\lambda_1\sigma_{y\eta} = \sigma_{y1}$ (3) $\sigma_{11} = \sigma_{11}$
(4) $\lambda_2\sigma_{y\eta} = \sigma_{y2}$ (5) $\sigma_{12} = \sigma_{12}$ (6) $\sigma_{22} = \sigma_{22}$
(7) $\sigma_{yz} = \sigma_{yz}$ (8) $\lambda_1\sigma_{\eta z} = \sigma_{1z}$ (9) $\lambda_2\sigma_{\eta z} = \sigma_{2z}$
(10) $\sigma_{zz} = \sigma_{zz}$

Identification rules

Identification of composite models is straightforward:[5]

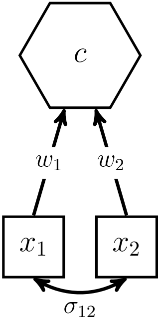
- Fix scale of all emergent variables, e.g., $\boldsymbol{w}_j'\boldsymbol{\Sigma}_{jj}\boldsymbol{w}_j = 1$
- Each emergent variable must be connected to at least one emergent variable or variable not forming the emergent variable

(► Structural model needs to be identified)

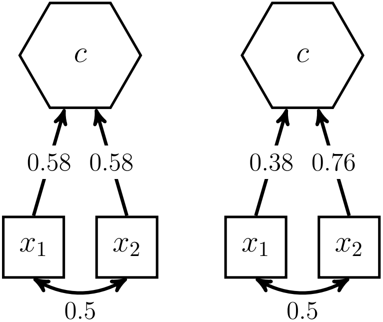
$\Rightarrow$ All model parameters can be uniquely retrieved from the population indicator covariance matrix

[5] Ignoring trivial regularity assumptions such as weight vectors consisting of zeros only; and similarly, we ignore cases where intra-block covariance matrices are singular.

Not identified composite model



$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$$



Infinite number of weight sets satisfying the scaling condition, i.e., variance of emergent variable equals 1: $\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w} = 1$.

Confirmatory composite analysis

Step 1: Model specification

Step 2: Model identification

Step 3: Model estimation

Step 4: Model assessment

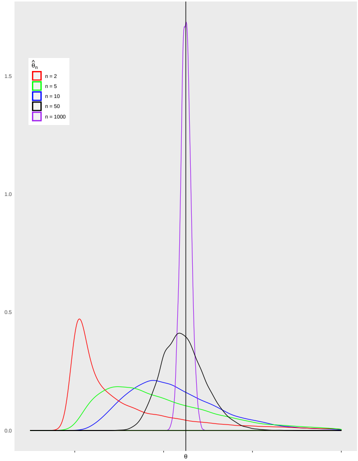
Model estimation

Typically the population parameters θ are unknown as we only have a sample at hand

$\Rightarrow$ Model parameters need to be estimated based on a sample (which is representative of the considered population)

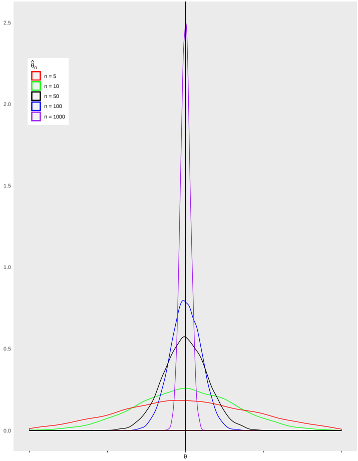
Which estimator to choose?

Consistent estimator



$$\text{plim}\,\hat{\theta} = \theta$$

Unbiased estimator



$$E(\hat{\theta}) = \theta$$

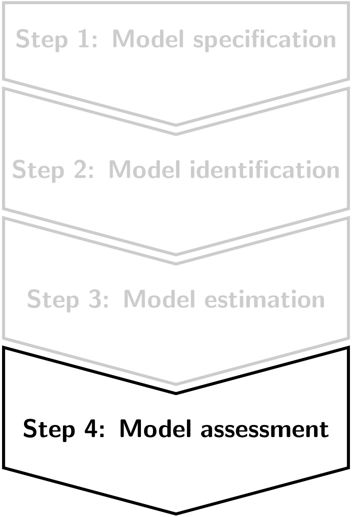
Potential Candidates

To estimate the parameters, several methods can be used:

- Approaches to generalized canonical correlation analysis (GCCA) such as MAXVAR (Kettenring, 1971)
- Regularized general canonical correlation analysis (RGCCA, Tenenhaus & Tenenhaus, 2011)
- Iterative partial least squares algorithm (PLS, Wold, 1975)
- Generalized structured component analysis (GSCA, Hwang & Takane, 2004)
- Maximum-likelihood estimator

Predetermined weights such as unit weights or expert weights are also conceivable.

Confirmatory composite analysis



Model assessment

Model assessment consists of two steps:

- Overall model fit assessment
- Assessment of the emergent variables

Overall model fit assessment

Overall model fit assessment is crucial to investigate whether the model is an acceptable description of reality.
The overall model fit can be assessed in two non-exclusive ways:

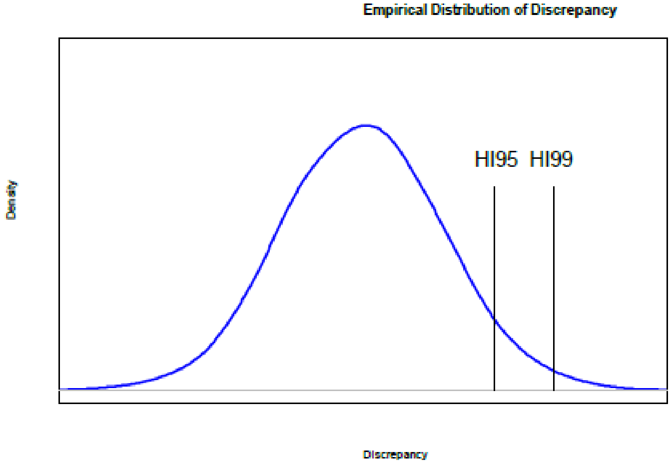
- Test for exact model fit
- Fit indices (heuristic rules)

Test for exact model fit

To test the exact model fit ($H_0 : \mathbf{\Sigma} = \mathbf{\Sigma}(\theta)$), a bootstrap-based test was proposed for CCA (Beran & Srivastava, 1985; Schuberth et al., 2018) in combination with various discrepancy measures such as

- Standardized root mean squared residual (SRMR)
- Geodesic distance (d_G)
- Squared Euclidean distance (d_L)

Bootstrap-based test for exact model fit



Fit indices

Tests for exact model fit are often regarded as too stringent.

The hypothesis of exact fit is often regarded as not plausible: "All models are wrong" (Box, 1976)

Approximate fit indices such as

- Standardized root mean squared residual (SRMR, Bentler, 1995; Schuberth et al., 2018), and
- Goodness of fit index (GFI, Cho et al., accepted; Jöreskog & Sörbom, 1989)

quantify the degree of misfit.

$\Rightarrow$ The larger their value, the larger the misfit.

Criticism on fit indices

Fit indices have been criticized:

- Fit indices are descriptive and non-inferential
- Derived thresholds are subjective and arbitrary

Assessment of emergent variables

Do the estimates align with your theory?

Compare

- estimated parameters,
- sign of the estimates, and
- significance of the estimate

with your expectations/theory

Assess potential multicollinearity issues, e.g., consider variance inflation factor.

Confusion

In 2020, CCA was dubbed as the method of confirming measurement quality (MCMQ) in PLS-SEM (Hair et al., 2020)

Problem

CCA and MCMQ are not the same!

Difference between CCA and MCMQ

Differences between CCA and MCMQ[6]

	Confirmatory composite Analysis	Method of confirming measurement quality
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.

$\Rightarrow$ Make sure to not confuse the CCA and MCMQ

[6] Schuberth, F. (2021). Confirmatory composite analysis using partial least squares: Setting the record straight. *Review of Managerial Science*, *15*(1).

1. History of Confirmatory Composite Analysis
2. Confirmatory Composite Analysis
   - Model Specification
   - Model Identification
   - Model Estimation
   - Model Assessment
3. Demonstration

Demonstration

Considering Value of Destination Experience (VDE)[7]:

"Also, there is one more second-order construct for the perceived value of destination **formed** from a scale measuring functional, social, and emotional value" (Lee et al., 2018, p.492)

Hypothesis:
VDE emerges from Functional, Social and Emotional Values within its environment.

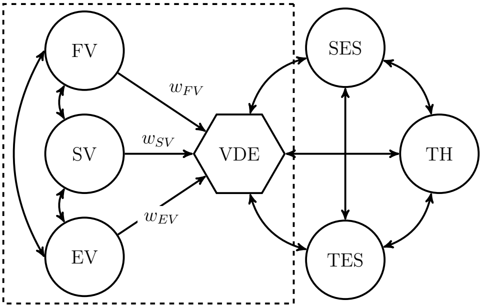
VDE's enivronment consists of:

- Service Experience Satisfaction (SES)
- Travel Experience Satisfaction (TES)
- Tourist Happiness (TH)

[7] Lee, H., Lee, J., Chung, N., & Koo, C. (2018). Tourists' happiness: Are there smart tourism technology effects? *Asia Pacific Journal of Tourism Research*, *23*(5), 486–501.

Demonstration

Model specification:



Model identification:

- Scale of the emergent variable is fixed (ensured by PLS)
- Emergent variable is not isolated

$\Rightarrow$ Model is identified

Demonstration

To conduct the analysis, ADANCO can be used:
https://www.composite-modeling.com/

Similarly, the R package cSEM can be used:
https://github.com/M-E-Rademaker/cSEM

References I

Bentler, P. M. (1995). *EQS structural equations program manual* (Vol. 6). Encino, CA, Multivariate Software, Inc.

Beran, R., & Srivastava, M. S. (1985). Bootstrap tests and confidence regions for functions of a covariance matrix. *The Annals of Statistics*, *13*(1), 95–115.

Box, G. E. P. (1976). Science and statistics. *Journal of the American Statistical Association*, *71*(356), 791–799.

Cho, G., Hwang, H., Sarstedt, M., & M., R. C. (accepted). Cutoff criteria for overall model fit indexes in generalized structured component analysis. *Journal of Marketing Analytics*.

Hair, 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.

References II

Henseler, J. (in press). *Composite-based structural equation modeling: Analyzing latent and emergent variables*. New York, The Guilford Press.

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

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

Hwang, H., & Takane, Y. (2004). Generalized structured component analysis. *Psychometrika*, *69*(1), 81–99.

References III

Jöreskog, K. G., & Sörbom, D. (1989). *LISREL 7 user's reference guide*. Scientific Software.

Kettenring, J. R. (1971). Canonical analysis of several sets of variables. *Biometrika*, *58*(3), 433–451.

Lee, H., Lee, J., Chung, N., & Koo, C. (2018). Tourists' happiness: Are there smart tourism technology effects? *Asia Pacific Journal of Tourism Research*, *23*(5), 486–501.

Schuberth, F. (2021). Confirmatory composite analysis using partial least squares: Setting the record straight. *Review of Managerial Science*, *15*(1).

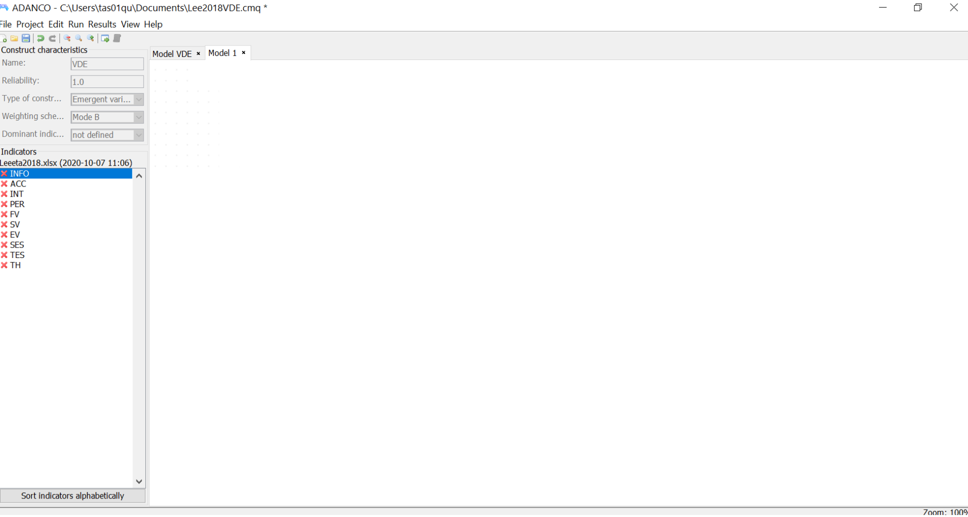
Schuberth, F., Henseler, J., & Dijkstra, T. K. (2018). Confirmatory composite analysis. *Frontiers in Psychology*, *9*(2541).

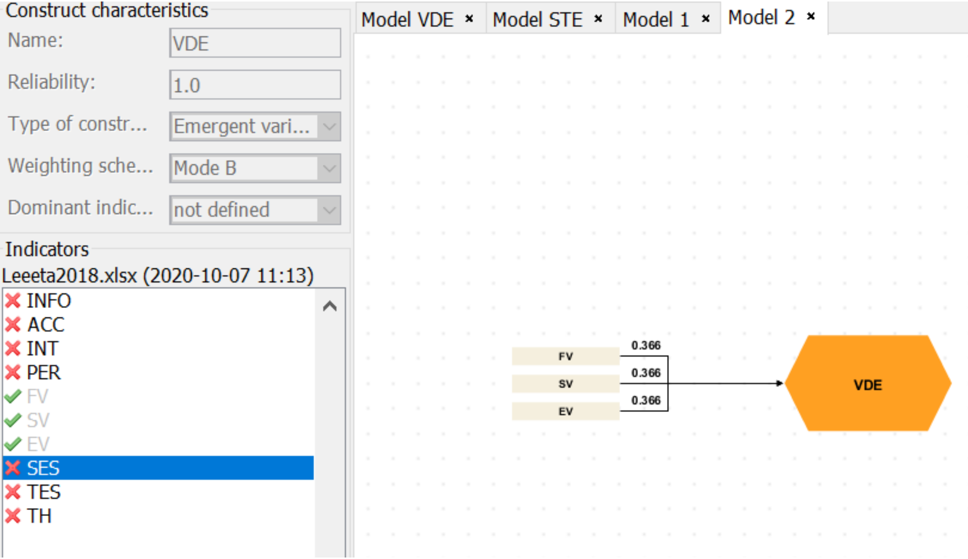
Tenenhaus, A., & Tenenhaus, M. (2011). Regularized generalized canonical correlation analysis. *Psychometrika*, *76*(2), 257–284.

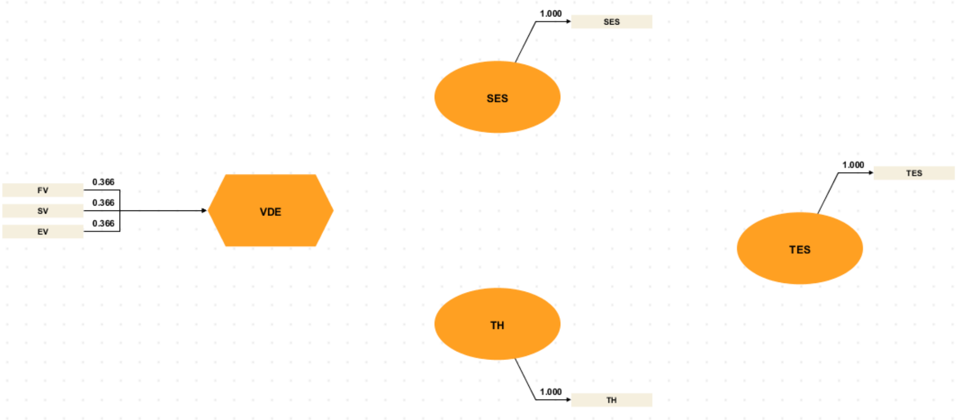
References IV

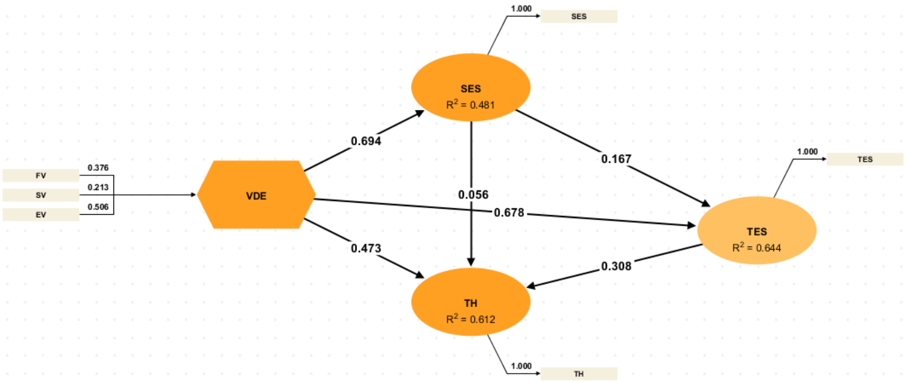
Wold, H. (1975). Path models with latent variables: The NIPALS approach. In H. Blalock, A. Aganbegian, F. Borodkin, R. Boudon, & V. Capecchi (Eds.), *Quantitative sociology* (pp. 307–357). New York, Academic Press.

Appendix
Steps in ADANCO

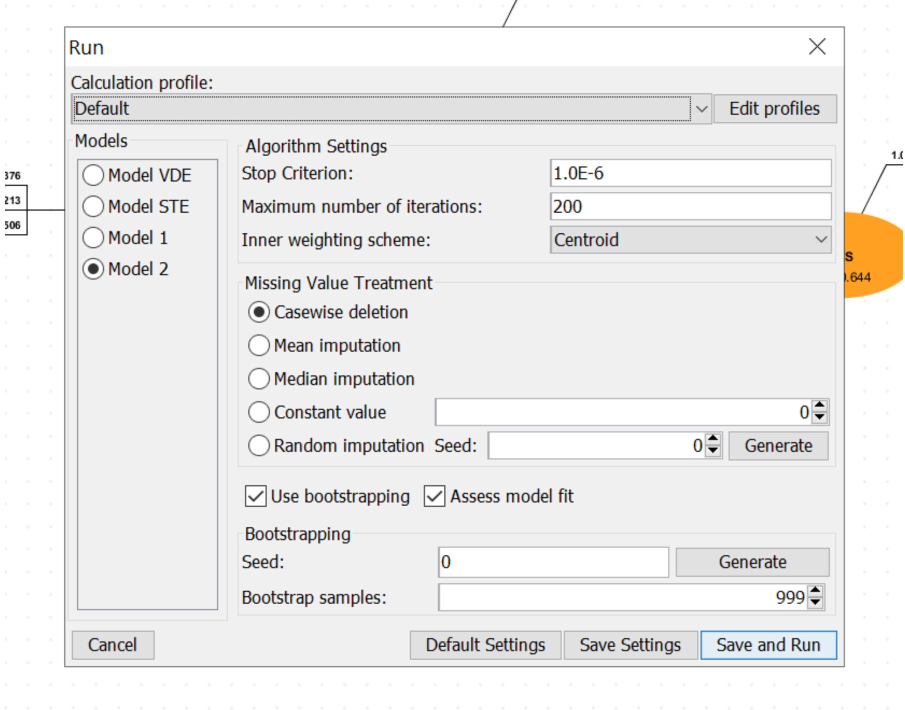

ADANCO - C:\Users\tas01qu\Documents\Lee2018VDE.cmq *
File Project Edit Run Results View Help
Construct characteristics
Name: VDE
Reliability: 1.0
Type of constr... Emergent vari...
Weighting sche... Mode B
Dominant indic... not defined
Indicators
Leeeta2018.xlsx (2020-10-07 11:06)
INFO
ACC
INT
PER
FV
SV
EV
SES
TES
TH
Sort indicators alphabetically
Model VDE
Model 1
Zoom: 100%

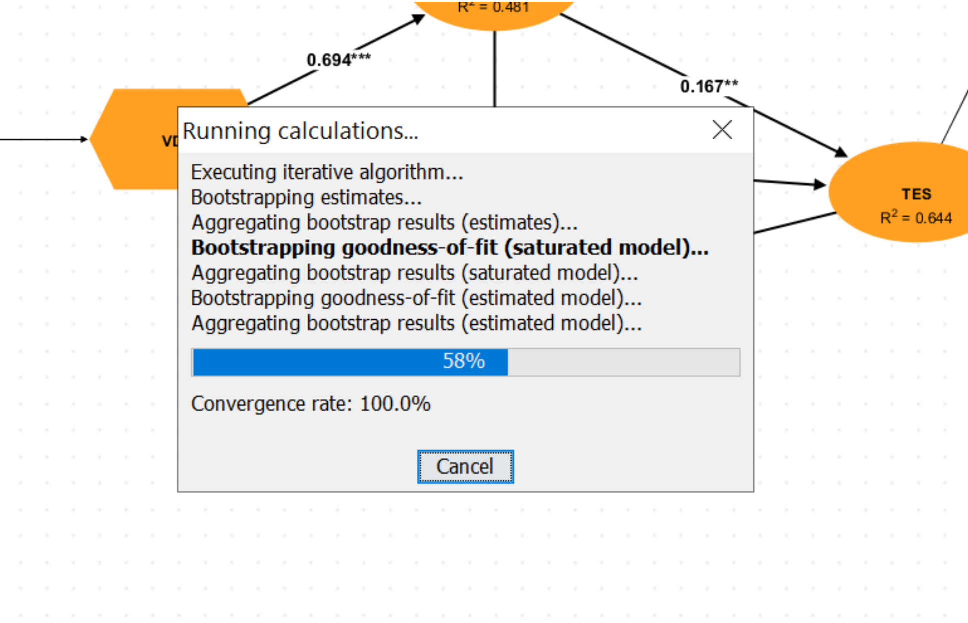

Construct characteristics
Name:
VDE
Reliability:
1.0
Type of constr...
Emergent vari...
Weighting sche...
Mode B
Dominant indic...
not defined
Indicators
Leeeta2018.xlsx (2020-10-07 11:13)
INFO
ACC
INT
PER
FV
SV
EV
SES
TES
TH
Model VDE
Model STE
Model 1
Model 2
FV
SV
EV
0.366
0.366
0.366
VDE

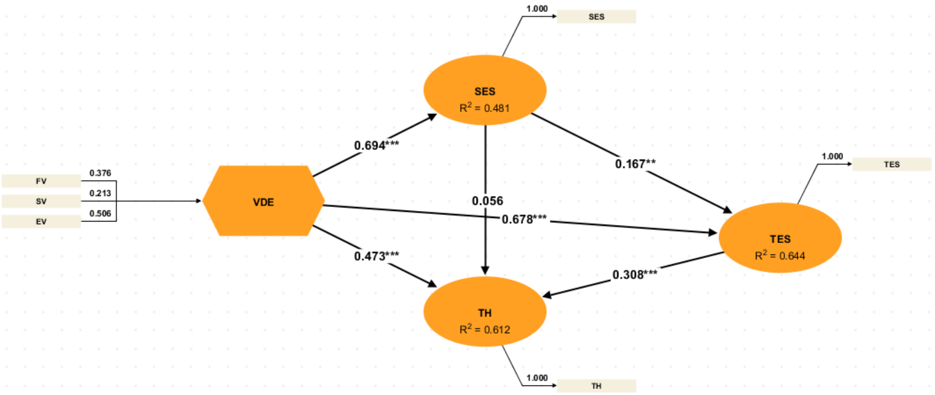

1.000
SES
SES
1.000
TES
FV
0.366
SV
0.366
EV
0.366
VDE
TES
TH
1.000
TH



1.000
SES
SES
R^2 = 0.481
0.694
0.167
1.000
TES
FV
0.376
SV
0.213
EV
0.506
VDE
0.056
0.678
TES
R^2 = 0.644
0.473
0.308
TH
R^2 = 0.612
1.000
TH


Run
Calculation profile:
Default
Edit profiles
Models
Model VDE
Model STE
Model 1
Model 2
Algorithm Settings
Stop Criterion:
1.0E-6
Maximum number of iterations:
200
Inner weighting scheme:
Centroid
Missing Value Treatment
Casewise deletion
Mean imputation
Median imputation
Constant value
0
Random imputation
Seed:
0
Generate
Use bootstrapping
Assess model fit
Bootstrapping
Seed:
0
Generate
Bootstrap samples:
999
Cancel
Default Settings
Save Settings
Save and Run


R² = 0.481
0.694***
0.167**
TES
R² = 0.644
Running calculations...
Executing iterative algorithm...
Bootstrapping estimates...
Aggregating bootstrap results (estimates)...
Bootstrapping goodness-of-fit (saturated model)...
Aggregating bootstrap results (saturated model)...
Bootstrapping goodness-of-fit (estimated model)...
Aggregating bootstrap results (estimated model)...
58%
Convergence rate: 100.0%
Cancel


1.000
SES
SES
R2 = 0.481
0.694***
0.167**
1.000
TES
FV
0.376
SV
0.213
EV
0.506
VDE
0.056
0.678***
TES
R2 = 0.644
0.473***
0.308***
TH
R2 = 0.612
1.000
TH