

Applied Psychometrics: The Application of CFA to Multitrait-Multimethod Matrices (CFA-MTMM)

Theodoros A. Kyriazos

Department of Psychology, Panteion University, Athens, Greece
Email: th.kyriazos@gmail.com

How to cite this paper: Kyriazos, T. A. (2018). Applied Psychometrics: The Application of CFA to Multitrait-Multimethod Matrices (CFA-MTMM). *Psychology*, 9, 2625-2648.
<https://doi.org/10.4236/psych.2018.912150>

Received: October 4, 2018

Accepted: November 13, 2018

Published: November 16, 2018

Copyright © 2018 by author and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The application of CFA to Multitrait-Multimethod Matrices (MTMM) is an elaborated method for the evaluation of construct validity in terms of the discriminant and convergent validity as well as method effects. It is implemented to evaluate the acceptability of a questionnaire as a measure of constructs or latent variables, during the structural and external stages of an instrument validation process. CFA can be carried out in MTMM matrices, like in any other covariance matrix, to study the latent variables of traits and methods factors. CFA-MTMM models can distinguish systematic trait or method effects from unsystematic measurement error variance, thus offering the possibility of hypotheses testing on the measurement model, while controlling possible effects. A plethora of CFA-MTMM approaches have been proposed, but the Correlated Methods Models and the Correlated Uniqueness Models are two CFA-MTMM methods most widely used. This work discusses how these two approaches can be parametrized, and how inferences about construct validity and method effects can be drawn on the matrix level and on parameters level. Then, the two methods are briefly compared. Additional CFA MTMM parameterizations are also discussed and their advantages and disadvantages are summarized.

Keywords

CFA MTMM, Construct Validity, Discriminant Validity, Convergent Validity, Method Effects, Correlated Methods Models, Correlated Uniqueness Models

1. Introduction

Test developers ought to publish only measurement instruments meeting the

guidelines specified by the APA standards (Franzen, 2002). Inaccurate measurements may lead to incorrect diagnoses, suboptimal treatment decisions or biased treatment effects estimates (Koch, Eid, & Lochner, 2018; Courvoisier, Nussbeck, Eid, Geiser, & Cole, 2008). Additionally, the adequacy of inferences made based on test scores is a central issue in social and behavioral sciences and it is related to validity (Messick, 1980, 1989, 1995; APA/AERA & NCME, 1999, 2014). Construct validity is the extent to which a measuring tool “reacts” the way the construct purports to react when compared with other, well-known measures of different constructs (DeVellis, 2017) and it consists the central focus of each measurement process (Kline, 2009) and an all-embracing principle of validity (Messick, 1995; Brown, 2015). Construct validity (Cronbach & Meehl, 1955) examines the theoretical relationship of a variable (like the scale score) to other variables (Kyriazos, 2018) and it incorporates the internal scale structure (Zinbarg, Yovel, Revelle, & McDonald, 2006; Revelle, 2018) or the correct measurement of variables intended to be examined (Kline, 2009). Construct validation is an ongoing process (Franzen, 2002).

To this end, Campbell and Fiske (1959) introduced the Multitrait-Multimethod (MTMM) Matrix, an evaluation method of the acceptability of questionnaires as measures of constructs or latent variables (Price, 2017). This method involves a correlation matrix customized to enhance the evaluation of construct validity¹ in terms of the discriminant and convergent validity (Brown, 2015). The method that Campbell and Fiske (1959) operationalize is based on the idea that theoretically related scores of the same or similar traits should also show high positive correlations with each other whereas theoretically unrelated scores of different constructs should also show low or insignificant correlations. The former refers to convergent validity and the latter refers to discriminant validity (Koch, Eid, & Lochner, 2018; Franzen, 2002).

The purpose of this study is to describe the application of Factor Analysis—EFA and mainly CFA—into the Multitrait-Multimethod Matrices (MTMM), as an elaborated approach of studying convergent validity, discriminant validity, and method effects during the estimation of construct validity.

2. Basic Concepts & Overview of the “Classic” MTMM

The Multitrait-multimethod (Campbell & Fiske, 1959) is “an analytic method that includes evaluation of construct validity relative to multiple examinee traits in relation to multiple (different) methods for measuring such traits” (Price, 2017: p. 140). MTMM arranges the analysis of validity evidence in a manner that transcends the common “set of associations” approach, permitting the examination of associations among several psychological variables, each being measured through several methods (Furr, 2011). More specifically, two or more traits are measured with two or more methods. Traits could be hypothetical constructs about stable characteristics, like cognitive abilities. On the other hand, methods

¹Construct Validity can be assessed by other methods along with MTMM. See Kyriazos et al. (2018a, 2018b, 2018c, 2018d, 2018e) for applied examples of construct validation studies.

can encompass different occasions, or means of data collection (e.g., self-report, structured interview and observation), or multiple-test forms, e.g. for different informants (teachers, parents and peers). Note that, trait interpretation suggests that a latent attribute or construct explains the consist patterns observed in the scores (Price, 2017; Kline, 2016). The MTMM method is carried out during the structural and external stages of the validation process (Price, 2017), see **Table 1** and **Figure 1** for an overview.

For example, during the development of a new self-report scale on social anxiety, the researcher could also measure three constructs (e.g., affect, self-esteem and depression) by three methods for each construct (i.e. self-report, report by partner, and report by peer). Correlations among all measures for all domains are calculated, and they are arranged in a “multitrait-multimethod” correlation matrix (Furr, 2011), like the one presented in **Table 2**.

For the simplest case of MTMM, at least two methods and two constructs are required (2 methods \times 2 traits) to generate a multitrait-multimethod design. Each score is a *trait-method unit* or TMU (Campbell & Fiske, 1959: p. 81; Koch et al., 2018). The MTMM correlation matrix (see **Table 2**) consists of four sets of correlations: 1) the monotrait-monomethod block correlations; 2) the monotrait-heteromethod block correlations; 3) the heterotrait-monomethod block correlations; and 4) the heterotrait-heteromethod block correlations (Koch et al., 2018; Furr, 2011). The correlations are divided into the two different method blocks of the table: 1) The *monomethod blocks*; and 2) the *heteromethod blocks* (Trochim, 2002; Brown, 2015). Additionally, two different sets of triangles are nested in the two blocks (Campbell & Fiske, 1959: p. 82; Trochim, 2002): 1) The heterotrait-monomethod triangles; and 2) The heterotrait-heteromethod triangles (**Table 2**). The general layout of an MTMM matrix is presented in **Table 2**. This example layout involves three different traits, each measured by three methods, generating nine separate variables. *The interpretation of the coefficients contained in the body of the table, is included in the table note.*

The monomethod blocks emerge from correlations of measurements by the same method. Their number equals the methods used. In order to account for measurement error (Koch et al., 2018), the reliability coefficients can be inserted

Table 1. Summary of the big picture of the place of MTMM in the construct validation process.

Stage MTMM is used	External stage/component
Purpose of MTMM	Study the external relationships of the target construct with other constructs and/or traits of the respondents
Question Asked	Do the target construct as measured by observed variables fit within the nomological network (Cronbach & Meehl, 1955) of expected construct relations?
Other Methods Used in Tandem	Group measurement invariance, Structural equation modeling, Correlation of construct with other constructs

Benson (1988) adapted by Price (2017) page 129.

Table 2. Sample MTMM Matrix of three Traits (Constructs) and 3 Methods.

		METHOD 1			METHOD 2			METHOD 3		
Traits		X1-1	X2-1	X3-1	X1-2	X2-2	X3-2	X1-3	X2-3	X3-3
METHOD 1	X1-1	(0.91)								
	X2-1	0.53	(0.91)							
	X3-1	0.4	0.39	(0.78)						
METHOD 2	X1-2	0.59	0.24	0.11	(0.95)					
	X2-2	0.24	0.59	0.12	0.7	(0.96)				
	X3-2	0.13	0.13	0.48	0.61	0.6	(0.86)			
METHOD 3	X1-3	0.58	0.24	0.13	0.69	0.44	0.35	(0.96)		
	X2-3	0.25	0.6	0.14	0.45	0.68	0.36	0.69	(0.94)	
	X3-3	0.13	0.13	0.47	0.36	0.34	0.6	0.6	0.62	(0.87)

Notes: Numbers in body of table are correlation coefficients except those in parentheses. Data is artificial. Cells in light Gray indicate Discriminant validity (different traits measured by same methods—should be lowest of all). Cells in bold typeface indicate Convergent validity (same trait measured by different methods—should be strong and positive; also termed Validity Diagonal). Cells in dark gray indicate Discriminant validity (different traits measured by different methods—should be lowest of all). Reliability coefficients for each test are in parentheses (Reliability Diagonal). Monomethod blocks are marked with a dotted line. Heteromethod blocks are marked by a thick line. The Table format was adopted by Price (2017: p. 133) and Brown (2015, p. 189).

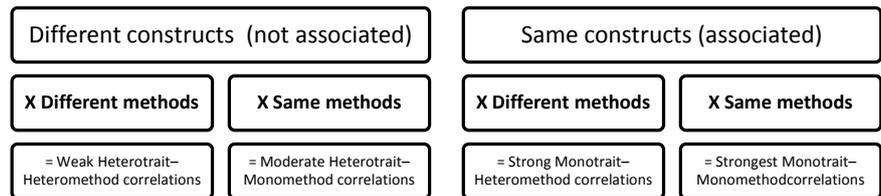


Figure 1. MTMM framework overview. Source: Content adapted by Furr (2011: p. 59).

in the diagonal of these blocks called Reliability Diagonal (Campbell & Fiske, 1959: p. 81). The reliability diagonal should preferably comprise the largest coefficients in the matrix (Campbell & Fiske, 1959; Price, 2017), i.e. the measure should show the strongest correlations with itself than with any other construct in the matrix (Trochim, 2002; Brown, 2015). For example, in Table 2 the reliability coefficients range from 0.78 to 0.96. Next, to the reliability diagonal, the off-diagonal elements are the *heterotrait-monomethod triangles* (Campbell & Fiske, 1959: p. 82; Trochim, 2002) presenting correlations evidencing the *discriminant validity* of different traits measured by same methods, thus they correlations should be the lowest of all (Price, 2017).

The *heteromethod blocks* emerge from correlations of measurements by different methods. Their number equals to $k(k - 1)/2$, where k is the number of methods used (Trochim, 2002). The correlations in their diagonals represent the convergent validity of the same trait measured by different methods. Thus, they should be strong and positive (Price, 2017; Brown, 2015). For example, in

Table 2 correlations suggesting convergent validity are ranging from $r = 0.47$ to $r = 0.69$. The off-diagonal elements in the heteromethod blocks called *heterotrait-heteromethod triangles* (Campbell & Fiske, 1959: p. 82; Trochim, 2002). They present correlations evidencing *discriminant validity* because different traits are measured by different methods. They also should be the lowest of all (Price, 2017). Brown (2015) explains that discriminant validity is demonstrated when weaker correlations emerge between different traits measured by different methods (i.e., heterotrait-heteromethod coefficients) in comparison to the correlations of the validity diagonal (i.e., that is monotrait-heteromethod coefficients). For example, in **Table 2**, discriminant validity is supported because the heteromethod block correlations in the off-diagonal elements are lower (from $r = 0.11$ to $r = 0.45$) than the validity coefficients (from $r = 0.47$ to 0.69). Finally, Brown (2015) explains that *method effects* may be evaluated by comparing the differential magnitude of correlations in the off-diagonal elements of the monomethod blocks.

2.1. Rules of Evaluation of the Data

Campbell and Fiske (1959: pp. 82-83) proposed four rules for the evaluation of MTMM matrices. The correlations in the monotrait-heteromethod block have to be statistically significant and strong. The correlations in the heterotrait-heteromethod block should be weak or should equal the correlations in the heterotrait-monomethod block. The correlations in the heterotrait-monomethod block should be smaller than the correlations in the monotrait-heteromethod block. The pattern of the trait intercorrelations in all heterotrait triangles should be similar for the mono- as well as the heteromethod blocks. This property also refers to the discriminant validity of the measures (summarized by Koch et al., 2018). For details refer also to Eid (2010). Finally, Trochim (2002) in line with Campbell and Fiske (1959) also suggests the following guidelines for the interpretation of an MTMM matrix: 1) Reliability coefficients must be the highest values in the matrix because the correlations between a trait and itself are expected to be the strongest than with any other construct (also Brown, 2015). 2) The coefficients presented in the validity diagonal must be statistically different from zero and preferably moderately strong. 3) Every coefficient in the validity diagonal must be the greatest column and row value of the method block it belongs. Recall that all off-diagonal elements must be the smallest to evidence discriminant validity. 4) Each validity coefficient must be greater than all the off-diagonal elements in the heterotrait monomethod triangles. If this is not true, then a method effect may be present. 5) The relationships of the different traits must be coherent in all triangles, i.e. relationships between every trait dyad must be cohesive, without areas of dramatic changes (Trochim, 2002). In spite of the ease of application of these rules, in practice, the application of MTMM method was criticized as having several limitations (Marsh, 1988, 1989; Schmitt & Stults, 1986; Eid, 2010).

2.2. Criticism of the “Classic” MTMM

Despite the fact that the [Campbell and Fiske \(1959\)](#) technique was a significant progress for examining construct validity, the MTMM approach was not widely accepted the years immediately following its inception ([Brown, 2015](#)). Additionally, several limitations were noted first by [Jackson \(1969\)](#). Jackson’s criticisms as summarized by [Franzen \(2002\)](#), mainly focused on the fact that the MTMM method juxtaposes distinct patterns without taking into account the overall structure of the relationships (i.e. between-network; [Cronbach & Meehl, 1955](#); [Martin & Marsh, 2006](#)). This argument is based on the idea that the correlation pattern may be affected by the way variance is distributed when measuring the target traits. To alleviate the MTMM disadvantages, [Jackson \(1969\)](#) suggested to carry out an EFA (PCA actually) in the monomethod matrix. Specifically, the correlation matrix is orthogonalized and then submitted to Principal Components Analysis, with a varimax rotation. The extracted factors equal the number of matrix traits. However, Jackson’s suggestion while offering some advantages, it assumes no relationships exist among the traits and this is rarely the case in psychology. Additionally, the influence of different measurement methods cannot be taken into account when the monomethod matrix is used ([Franzen, 2002](#)). Finally, EFA failed to obtain meaningful solutions of MTMM data due to the inherent restrictions of EFA to specify correlated errors ([Brown, 2015](#)).

Later, [Franzen \(2002\)](#) continues [Cole, Howard, and Maxwell \(1981\)](#) compared the MTMM results with constructs mono-operationalization suggesting that as a rule, the validity coefficients were misleadingly deflated. Therefore, the MTMM method and factor analysis may be used in tandem ([Franzen, 2002](#)). Finally, [Cole \(1987\)](#) repeated to a certain extent Jackson’s criticisms, i.e. absence of concrete guidelines regarding the size of the zero-order correlation coefficients. Additionally, Cole argued that the cost of the MTMM application in clinical settings deems the method inapplicable ([Franzen, 2002](#)). Crucially, Cole noted that the MTMM is prone to correlated errors due to factors like the time of day, or the instrument administrator. Cole suggested the use of CFA instead to examine discriminant and convergent validity ([Franzen, 2002](#)). For a detailed description of MTMM limitations see also [Marsh \(1988, 1989\)](#) and [Schmitt & Stults \(1986\)](#).

At the same time, non-factorial oriented attempts to find an appropriate method of statistical analysis for the data in the matrix were numerous ([Eid et al., 2008](#); [Eid, 2010](#)). These methods were mainly based on analysis of variance models or nonparametric alternatives and the reason for the lack of success was attributed to conflicting data in the diagonals in relation to the data in the triangles ([Sawilowsky, 2007](#)). For example, [Sawilowsky \(2002, 2007\)](#) proposed a statistical test for analysis MTMM data where the null hypothesis is that the coefficients in the matrix are unordered from high to low. This is tested against the alternative hypothesis of an increasing trend from the lowest level (heterotrait heteromethod) to the highest level i.e. reliability coefficients ([Sawilowsky, 2007](#):

p. 180). See Sawilowsky (2002, 2007) for more details and an applied example.

Despite the several alternatives proposed—like Sawilowsky's (2002, 2007)—the analysis of MTMM data within the CFA framework (a.k.a. CFA-MTMM models) has been a widely accepted methodological strategy (Eid et al., 2008; Byrne, 2012; Koch et al., 2018).

3. CFA Approaches to Analyzing the MTMM Matrix (CFA-MTMM Models)

Actually, MTMM matrix regained interest upon the realization that EFA procedures (Jackson, 1969) or CFA (Cole et al., 1981; Cole, 1987; Flamer, 1983; Marsh & Hocevar, 1983; Widaman, 1985) are equally applicable for MTMM data analysis. Specifically, CFA can be carried out in MTMM matrices, like in any other covariance matrix, to study the latent dimensions of traits and methods factors (Brown, 2015).

More specifically, CFA-MTMM models have many advantages as Koch et al., (2018) comment, because they offer a plethora of possibilities like: 1) distinguishing systematic trait or method effects from unsystematic measurement error variance (see Brown, 2015; Brown & Moore, 2012); 2) associating latent trait and/or method variables to other variables in order to study trait or method effects; and 3) testing specific hypotheses on the measurement model (Dumenci, 2000; Eid et al., 2006), while controlling possible effects (Franzen, 2002). Besides that, many different MTMM approaches have been suggested, such as variance component models or multiplicative correlation models (Koch et al., 2018 also quoting Browne, 1984; Dudgeon, 1994; Millsap, 1995a, 1995b; Wothke, 1995; Wothke & Browne, 1990).

Among them (see Marsh & Grayson, 1995; Widaman, 1985; Koch et al., 2018) two have been most widely used (Brown, 2015; Byrne, 2012): 1) the *Correlated Methods Models*, and 2) the *Correlated Uniqueness Models*.

3.1. Correlated Methods Models

This method was among the first significant developments in the field of CFA-MTMM models (Koch et al., 2018) and was developed by Keith Widaman (1985) proposing a nested model comparison.

The model comparison presented by Widaman (1985) had (t) traits and (m) methods where (t) is the number of trait factors and (m) is the number of method factors, abiding by five key specifications rules (Brown, 2015): 1) for model identification, at least three traits (t) and three methods (m) are required; 2) (t) \times (m) indicators are specified to define (t) + (m) factors; 3) each indicator loads on two factors, i.e. a trait factor and a method factor and all other cross-loadings are set to zero; 4) trait factors inter-correlations and method factors inter-correlations are freely estimated; 5) trait-method factors correlations are typically constrained to zero; and 6) indicator uniquenesses are estimated freely but cannot be specified to correlate with uniquenesses of other indicators. Each indicator in this specifi-

ication is a function of a trait, method, and unique factors, [Brown \(2015\)](#) explains.

The specified MTMM model becomes a baseline model, compared against a set of nested models in which parameters are either removed or added. Model comparison criteria used are the χ^2 ($\Delta\chi^2$) and other fit indicators (see [Byrne, 2012](#)), in a procedure similar to the multi-group CFA. Based on this comparison, evidence of convergent and discriminant validity emerges, both at the matrix level and at the individual parameter level ([Byrne, 2012](#)). [Figure 2](#) is a path diagram for an MTMM matrix of 3 traits and 3 methods to describe the steps of the correlated methods. Testing for evidence of convergent and discriminant validity involves comparisons between the model in [Figure 2](#) and three alternative nested MTMM models in [Figures 3-5](#) ([Byrne, 2012](#); [Brown, 2015](#)).

The Steps of the Correlated Methods

[Byrne \(2012\)](#) describes the steps for comparing the nested models as follows.

Step 1—The CTCM Model: Test the first model (see path diagram in [Figure 2](#)) is the hypothesized baseline model against which the three alternative CFA models are compared. A common problem in this step of the process is a non-positive matrix error (see [Kenny & Kashy, 1992](#); [Marsh, 1989](#)). Usually, a negative variance is associated, either with a residual or a factor and this problem is generally attributed to overparameterization of the MTMM models, due to their inherent complexity ([Byrne, 2012](#); quoting also [Wothke, 1993](#)). Possible

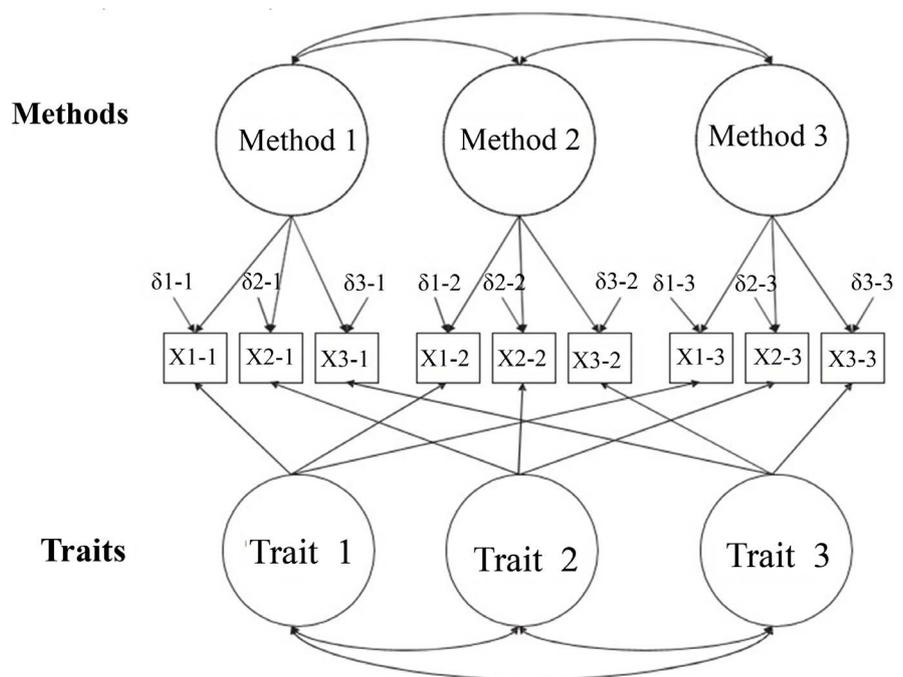


Figure 2. Model 1—The Correlated Traits/Correlated Methods Model (CTCM). Example of Correlated Methods CFA MTMM with 3 Traits and 3 Methods. Model 1 of 4 models that are specified (Baseline Model).

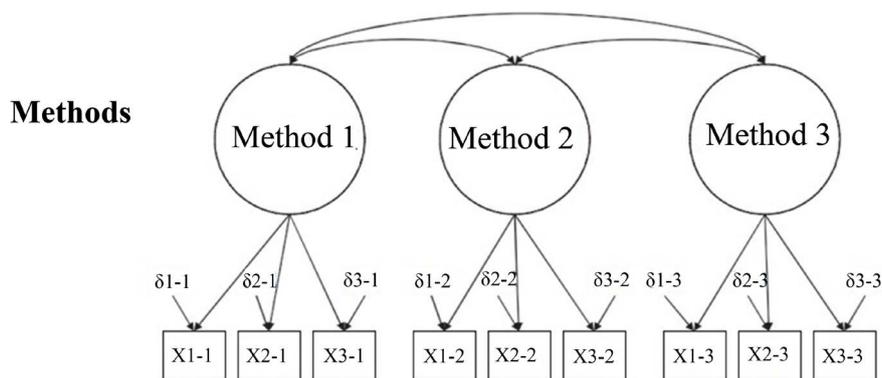


Figure 3. Model 2—The No Traits/ Correlated Methods Model (NTCM). Example of Correlated Methods CFA MTMM with 3 Traits and 3 Methods. Model 2 of 4 Nested models that are specified (Nested model to test Convergent Validity).

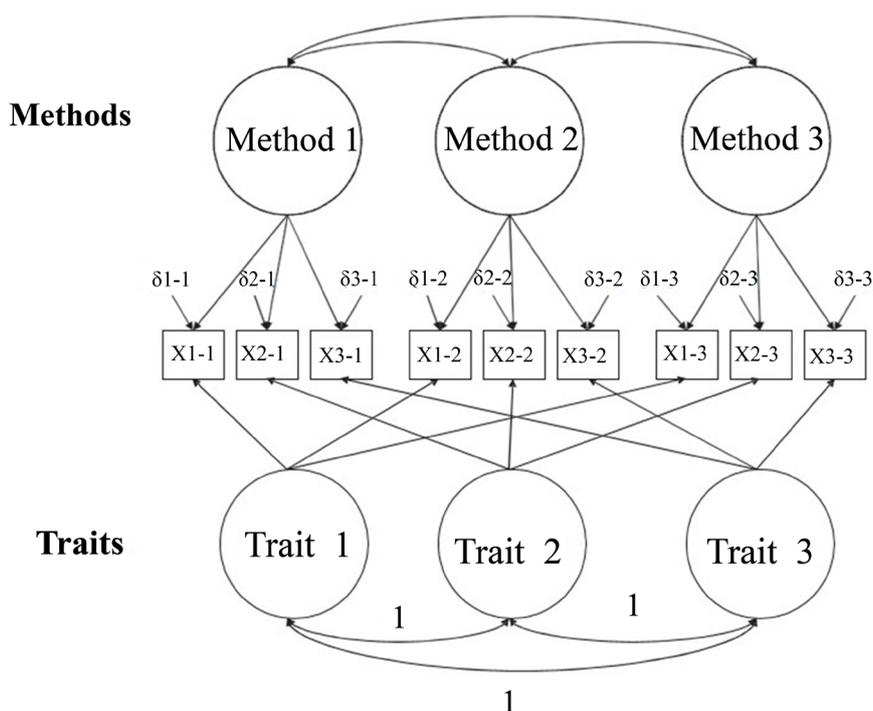


Figure 4. Model 3—Perfectly Correlated Traits/ Freely Correlated Methods Model (PCTCM). Example of Correlated Methods CFA MTMM with 3 Traits and 3 Methods. Model 3 of 4 Nested models that are specified (Nested model to test Discriminant Validity).

solutions are provided by Marsh et al. (1992) and Byrne (2012). Hopefully, upon model re-specification, the Correlated traits/Correlated Methods model (CTCM) results in a solution and the fit of this CTCM model must be considered (Byrne, 2012).

Step 2—The NTCM Model: The next hypothesized model of correlated methods specified is a no traits/correlated methods (NTCM) model (see path diagram in Figure 3). Its specification is identical to that of the CTCM model, except that there are no trait factors, as shown in Figure 3 (Widaman, 1985; Byrne,

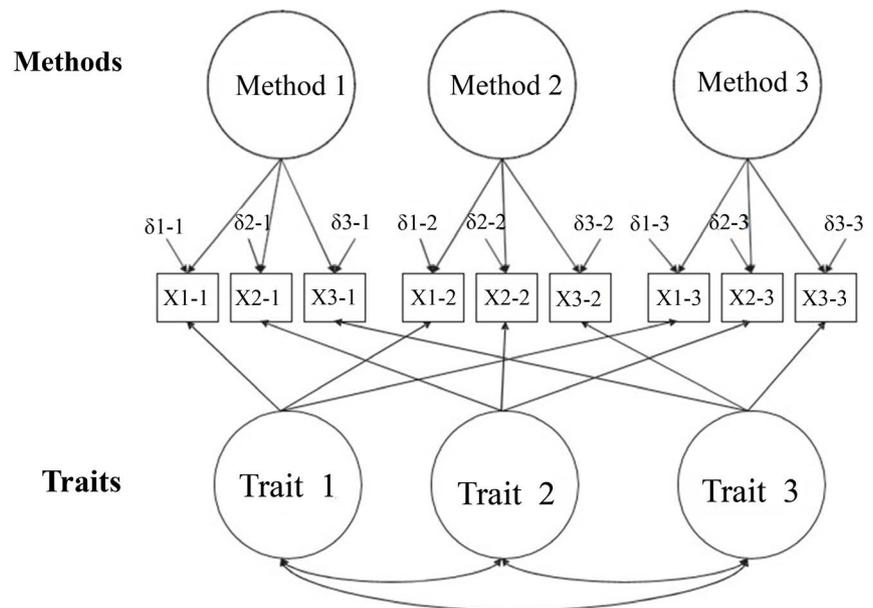


Figure 5. Freely Correlated Traits/Uncorrelated Methods Model (CTUM). Example of Correlated Methods CFA MTMM with 3 Traits and 3 Methods. Model 4 of 4 Nested models that are specified (Nested model to test Discriminant Validity).

2012). Because these two models are nested, a comparison of this model to the CTCM provides a statistical evaluation of whether effects associated with the different traits emerge (Brown, 2015). First, the goodness-of-fit for this NTCM model is examined in relation to the CTCM model.

Step 3—The PCTCM Model: The specification of a hypothesized Perfectly Correlated Traits/Freeely Correlated Methods (PCTCM) MTMM model follows. The path diagram of this model is presented in Figure 4. As with the hypothesized CTCM model (Figure 2), each observed variable loads on a trait and a method factor. Nevertheless, in this MTMM model, the traits are perfectly correlated (i.e., equal to 1.0). The method factors are freely estimated in this model, in line with both the CTCM model and NTCM model. The goodness-of-fit of this model is examined and it is compared to the fit of the previous models (in Figure 2 and Figure 3 respectively).

Step 4—The CTUM Model: The last correlated methods MTMM model is specified which is a Freely Correlated Traits/Uncorrelated Methods (CTUM) see Figure 5. It is different from Model 1 only in the absence of correlations in the method factors. The goodness of-fit for this model is evaluated (see also Byrne, 2012).

Step 5—Evaluation of Construct Validity at the Matrix Level: After the evaluation of goodness-of-fit of the four MTMM models and with the assumption that the fit is acceptable (within the acceptability criteria used for typical CFA models, e.g. Hu & Bentler, 1999; Brown, 2015), evidence of construct and discriminant validity is examined. This information at the matrix level derives through the comparison of particular pairs of models. Note that typically, a

summary of goodness-of-fit statistics of all four MTMM models is presented in a table, and a summary of model comparisons in a second table, see an example of both in **Table 3** adapted by *Byrne (2012)*.

Step 5a—Convergent Validity at the Matrix Level: Generally, the principles applied when testing for invariance are also applied when comparing models in this step (see **Table 3**). In line with the procedure proposed by *Widaman (1985)*, evidence of convergent validity is tested by comparing the CTCM model (Model in **Figure 2**) to the NTCM model (Model in **Figure 3**). When the difference in χ^2 ($\Delta\chi^2$) between these models is highly significant and the fit difference (ΔCFI) is also large evidence of convergent validity is suggested (*Byrne, 2012*).

Step 5b—Discriminant Validity at the Matrix Level: Discriminant validity is usually evaluated considering both traits and methods. Evidence of discriminant validity among traits emerges from the comparison of the CTCM Model in which traits are freely correlated (see Model in **Figure 2**) with the PCTCM model, specified in step 3 in which traits are perfectly correlated (see Model in **Figure 4**). The larger the difference between the values of χ^2 and CFI, the stronger the evidence of discriminant validity. The procedure is repeated to show evidence of discriminant validity regarding method effects. Specifically, the CTCM Model in which methods are freely correlated—specified in step 1 and presented as path diagram in **Figure 2**—, is compared to the CTUM model in which methods are uncorrelated-specified in step 4 and presented as path diagram in **Figure 5**. A large $\Delta\chi^2$ (and/or considerable ΔCFI) suggests common method bias, therefore the lack of discriminant validity across measurement methods (*Byrne, 2012*). For an applied example see *Byrne (2012)*. See Goodness of fit criteria used in **Table 3** and the Nested Models Compared in **Table 4**. *Byrne* suggests that the difference criteria for the comparison of the nested models to be identical to measurement Invariance criteria (*Cheung & Rensvold, 2002*) “as an evaluative base upon which to determine evidence of convergent and discriminant validity” (*Byrne, 2012: p. 300*). For a more detailed description and an applied example see *Byrne (2012)*.

Step 6—Evaluation of Construct Validity at the Parameter Level: Trait- and method-related variance can be examined more thoroughly by examining individual parameter estimates, i.e. the factor loadings and factor correlations of

Table 3. List of Goodness-of-Fit Statistics for CFA Multitrait-Multimethod (MTMM).

Goodness of fit Index for CFA MTMM Models 1 - 4 (Correlated Methods)	Models tested in Correlated Methods
Chi-square (χ^2)	
Degrees of Freedom	Model 1: CTCM Model (see Figure 2)
Comparative Fit Index (CFI)	Model 2: NTCM Model (see Figure 3)
Root Mean Square Error of Approximation (RMSEA)	Model 3: PCTCM Model (see Figure 4)
RMSEA 90% Confidence Interval (CI)	Model 4: CTUM Model (see Figure 5)
Standardized Root Mean Square Residual (SRMR)	

Notes: Based on a table used by *Byrne (2012: p. 299)*; Goodness of fit criteria are the same as standard CFA.

Table 4. Correlated methods Comparison of Nested Models in CFA MTMM (Correlated methods).

CFA MTMM Models 1 - 4 (Correlated Methods)	Differential Goodness-of-Fit
Model Comparisons	Difference in
Test of Convergent Validity	Chi-square (χ^2)
Model 1 (CTCM) versus Model 2 (NTCM)	Degrees of Freedom
Test of Discriminant Validity	Comparative Fit Index (CFI)
Model 1 (CTCM) versus Model 3 (PCTCM)	Significance level (p value)
Model 1 (CTCM) versus Model 4 (CTUM)	

Notes: Based on a table used by Byrne (2012: p. 300); Byrne suggests difference criteria adopted to be identical to measurement Invariance Criteria (see Cheung and Rensvold, 2002; Chen, 2007 for cutoff values).

the CTCM model (the Model of step 1, see path diagram in **Figure 2**) as Byrne (2012) explains.

Step 6a—Convergent Validity at the Parameter Level: The size of the trait loadings indicates convergent validity. First, all loadings are examined to see if they are statistically significant. Then all factor loadings across traits and factor loadings across methods are compared, to see whether method variance or trait variance is higher over all the available ratings (Byrne, 2012).

Step 6b—Discriminant Validity at the Parameter Level: Factor correlation matrices contain the evidence of discriminant validity regarding particular traits and methods. Generally, correlations among traits should theoretically be small to show evidence of discriminant validity, but such findings are rare in psychology. Regarding the correlations of the method factors, discriminability of these parameters values suggests the extent to which the methods are dissimilar, testing a major assumption of the MTMM strategy (Campbell & Fiske, 1959; Byrne, 2012). For an applied example of the correlated method refer to Byrne (2012) and Brown (2015).

Some simpler alternatives to CTCM models have been proposed, and the correlated-uniqueness (CU) model described next is one of them (Kline, 2016).

3.2. Correlated Uniqueness Models (CU)

The CU models signify a special case of the general CFA model based on the work of Kenny (1976, 1979) and Marsh (1988, 1989) who proposed this alternative MTMM model parametrization to avoid the common estimation and convergence problems arising when analyzing CTCM models, like the one in **Figure 2**, (Brown, 2015; Byrne, 2012; also quoting Kenny & Kashy, 1992; Marsh & Bailey, 1991; Tomás, Hontangas, & Oliver, 2000).

The CU model identification must abide by the following rules as reported by Brown (2015). There must be at least two traits (t) and three methods (m). However, a $2 t \times 2 m$ model can be also specified in the factor loadings of indicators that belong to the same trait factor are fixed to be equal. Rules regarding

the CU trait model are the same as in the correlated methods approach. Specifically: 1) Each indicator must load on one trait factor with all other cross-loadings fixed to zero; and 2) The trait factor correlations are freely estimated. Consequently, the major difference between the CU method in comparison to the correlated methods is the estimation of method effects. In the CU model, method effects are assessed by adding correlated uniquenesses (errors) among the indicators of the same assessment method instead of specifying method factors (Brown, 2015). **Figure 6** presents the path diagram of the correlated uniqueness CFA specification for an MTMM matrix with 3 traits and 3 methods.

After the model specification, the model is tested and the fit of the solution to the data is evaluated. Except fit the inspection of standardized residuals, the modification indices must be also evaluated (as proposed by Brown, 2015). The parameter estimates Brown (2015) continues are then examined. That is, when loadings of traits factors are substantial and statistically significant, this is an evidence supporting convergent validity. Conversely, when correlations among the trait factors are substantial, this is indicative of low discriminant validity. Method effects are assessed when correlated uniquenesses among indicators of the same method are moderate or large in magnitude (Brown, 2015).

For example, Brown (2015) explains, when trait factor loadings are large this suggests that the relation of the indicators to their purported latent constructs is strong thus, convergent validity can be supported while adjusting for method effects. When modest correlations emerge among trait factors this is indicative of adequate discriminant validity. The significance of method effects is evaluated by the size of the correlated uniquenesses among indicators. As in the classic MTMM matrix, the method effects in the CFA model must be smaller than those of the indicators. For an applied example of the CU, method sees Byrne (2012)

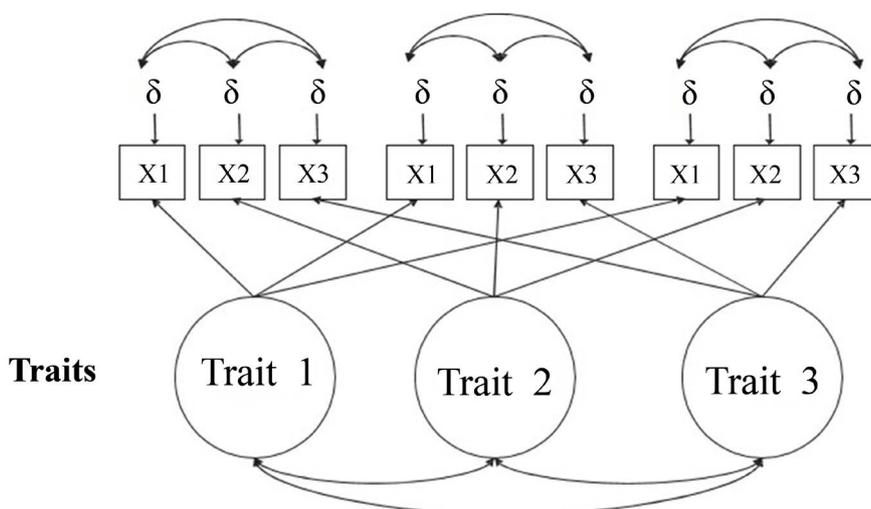


Figure 6. The Correlated Uniquenesses Model (CU). Path diagram of the correlated uniquenesses CFA specification for an MTMM matrix of 3 traits and 3 methods. Method effects are evaluated by specifying correlated uniquenesses (deltas or δ) among the indicators of the same assessment method.

and Brown (2015). Saris and Aalberts (2003) also examined different interpretations of CU models.

3.3. Comparison of the Two Methods

Some recommend the correlated uniquenesses (CU) method (Kenny, 1976, 1979; Kenny & Kashy, 1992; Marsh, 1989) as superior, whereas others support the general CFA method (Conway, Scullen, Lievens, & Lance, 2004; Lance, Noble, & Scullen, 2002). Nonetheless, Byrne (2012) comments that the general CFA model is as a rule preferable.

Specifically, an advantage of the CT-CM model is that it permits fragmenting observed variance into a trait, method, and error variance in comparison to the CU model that is more restrictive (Eid et al., 2006; Geiser et al., 2008; Koch et al., 2018). Kenny and Kashy (1992) argue that this fragmentation of the correlated methods model is the reason why this method corresponds directly to Campbell and Fiske's (1959) original conceptualization of the MTMM matrix. Moreover, the parameter estimates of the correlated methods models offer direct evidence of construct validity. For instance, substantial trait factor loadings show convergent validity; small and/or non-significant method factor loadings suggest that method effects are absent and modest trait factor correlations suggest good discriminant validity. The addition of method factors permits interpretation of method effects (Brown, 2015).

Nevertheless, a drawback of the correlated methods model is that it is prone to underidentification errors, or Heywood cases and large standard errors. Because of these problems with the correlated methods model, methodologists suggested (Kenny & Kashy, 1992; Marsh & Grayson, 1995) using the correlated uniqueness model for the analysis of MTMM data. On the other hand, others (Lance, Noble, & Scullen, 2002) argue that given the substantive strengths of the correlated methods model, the CU model should be preferred only as last resort, i.e. when the correlated methods model fails to be identified (Brown, 2015).

For a summary of CFA-MTMM models see Dumenci (2000), Eid, Lischetzke, and Nussbeck (2006), as well as Shrout and Fiske (1995) and Koch et al. (2018). For applied examples of the above-mentioned procedure both for the Correlated methods and the Correlated Uniquenesses you can refer to Byrne and Goffin (1993), Byrne and Bazana (1996) as well as to Meyer, Frost, Brown, Steketee, and Tolin (2013) or Koch et al. (2018). Byrne (2012) refers readers who wish to read more on comparisons of the correlated uniquenesses, composite direct product, and general CFA models to other sources (Bagozzi, 1993; Bagozzi & Yi, 1990; Coenders & Saris, 2000; Hernández & González-Romá, 2002; Marsh & Bailey, 1991; Marsh, Byrne, & Craven, 1992).

3.4. Other CFA Parameterizations of MTMM Data

More recent, Koch et al. (2018) comment, MTMM research has gained insight on many aspects of the CTCM minus one (CT-C[M-1]) model coined by Eid,

2000; Eid, Lischetzke, Nussbeck, & Trierweiler, 2003; see also Geiser, Eid, & Nussbeck, 2008; Maydeu-Olivares & Coffman, 2006; Pohl & Steyer, 2010. These comprise the inspection of specifically correlated residuals (Cole, Ciesla, & Steiger, 2007; Saris & Aalberts, 2003), and extension of the method to longitudinal data (Courvoisier et al. 2008; Grimm, Pianta, & Konold, 2009; LaGrange & Cole, 2008) as well as to multilevel data (Hox & Kleiboeer, 2007).

More specifically, the CT-C (M-1) model is analogous to the correlated methods model (presented in **Figure 2**) but it comprises one method factor less than the traits included (M-1). For example, the model presented in **Figure 2** could be transformed into a CT-C (M-1) model by eliminating the Method 1 latent variable. This specification answers to identification problems of the correlated methods model (Brown, 2015). They can be defined in three steps: 1) Selection of a reference method (standard), on the basis of theory and research questions (Geiser, Eid, & Nussbeck, 2008). 2) Selection of general trait variables (common or indicator-specific) and M-1 method factors. 3) Define additional constraints as a function of how non-correlationally measurement errors are, how non-correlationally method variables with trait variables of the same TMU are, and what is the homogeneity of the variables. Note, that if parameters of the model are restricted, the model data is equivalent to the common factor TMU model (Koch et al., 2018 also quoting Geiser et al., 2008; Geiser, Eid, West, Lischetzke, & Nussbeck, 2012).

Latent difference models, Koch et al. (2018) argue, offer a possibility to contrast methods. Initially, they were developed for longitudinal research (Steyer, Partchev, & Shanahan, 2000) but they can successfully use in multimethod research (see Pohl, Steyer & Kraus, 2008). This technique also involves a reference method. Nevertheless, the latent method variables are not specified as residuals regarding the true-score variables of the reference method. It is hypothesized that each indicator taps a trait variable, but method variables are unidimensional. Therefore, a common method factor is specified for all indicators of the same method and the same trait (Koch et al., 2018).

One more alternative approach is the *direct product model* (Browne, 1984; Cudeck, 1988; Wothke & Browne, 1990; Verhees & Wansbeek, 1990). In this approach, the method factors relate to the trait factors in a multiplicatively instead of additively. That is, method effects may increase the correlations of strongly correlated traits, more than they increase the correlations of weakly correlated traits. In other words, the higher the correlation between the traits, the greater the method effects. However, the data must be suitable for the direct product model parameterization. Additionally, the implementation and interpretation of the method are rather complex and prone to improper solutions, but less prone than the correlated methods (Brown, 2015; Lievens & Conway, 2001; Campbell & O'Connell, 1967).

On the other hand, methodologists disagree on the merits of the multiplicative versus the additive models. Some support the additive ones as more successful

(Bagozzi & Yi, 1990), whereas others either disagreed (Goffin & Jackson, 1992; Coovert et al., 1997) or present mixed evidence (Byrne & Goffin, 1993). For example, Corten et al. (2002) reported better results in the additive models in 71 out of 79 datasets used (Loehlin & Beaujean, 2017; Loehlin, 2004).

CFA-MTMM models have received additional criticism because the meaning of the trait and method factors often remains unclear. This happens because CFA-MTMM analysis just assumes that trait and method factors are present. In order to properly interpret trait and method effects, what the latent factors mean must be clear (Koch et al., 2018). Koch et al., argue that developments in stochastic measurement theory (see Steyer, 1989; Zimmerman, 1975) proposed a technique to define the MTMM model factors in terms of random variables in a well-defined random experiment. Koch et al. (2018) add that a wide range of models for different purposes and data structures can be defined with this method but not all the CFA MTMM models.

With so many alternative approaches, is difficult for the researcher to choose the most appropriate. One of the most important criterion for selecting the appropriate model is the type of method to use. The researcher must examine: 1) the methods used in the study; 2) the data and the structure of the random experiment implied by the measurement design, and 3) the possible conceptualization of method effects. For a study with interchangeable methods, a classic CFA modelling strategy can be chosen. For a study with structurally different methods, several other modelling alternatives can be used. Each model has distinct features that should be appropriate for the research question of the study. This means, there is no optimal model for all types of CFA MTMM measurement designs. The selection of the model must be driven by theory and hypotheses studied (Koch et al., 2018). For comparisons of these alternatives, readers can refer to other sources (Eid et al., 2008; Pohl & Steyer, 2010; Saris & Aalberts, 2003).

Regarding CFA-MTMM research in general, Loehlin and Beaujean (2017) and Loehlin (2004) review that the implementation of CFA MTMM models within and across groups has also been discussed (Marsh & Byrne, 1993). Marsh and Hocevar, (1988) reported using CFA MTMM with higher-order structures (Loehlin & Beaujean, 2017; Loehlin, 2004). For applied examples the reader can refer to Bagozzi and Yi (1990), Coovert, Craiger, and Teachout, (1997) and Lievens and Conway (2001). Next we will focus on the longitudinal CFA-MTMM research.

4. MTMM Analysis in Longitudinal Research

Generally, longitudinal measurement designs permit to focus on: 1) change and/or variability over time; 2) measurement invariance (MI), or indicator-specific effects; and 3) examine potential causal relationships (Coulacoglou & Saklofske, 2017 also quoting Steyer, 2005).

Longitudinal CFA-MTMM models (or multitrait-multimethod-multioccasion; MTMM-MO) present numerous advantages. For example, they permit researchers

to study the true stability (i.e., without measurement error) and the true change of the target constructs in tandem with method effects inherent to different measurement levels (for example rater and target-construct level). In other words, researchers can focus on the change of the target construct over time (e.g. positive affect) taking also into account the change of different method effects over time. Furthermore, each measurement occurrence can be analyzed with regards to the convergent and discriminant validity of measures used. The concept of change of convergent validity over time may be particularly pertinent in intervention studies. Furthermore, the effect of external covariates can be regressed into the model like gender, age, SES or personality variables (Koch et al., 2018).

Koch, Schultze, Eid, and Geiser (2014) introduced a longitudinal MTMM modeling frame for numerous MTMM methods termed Latent State-Combination-Of-Methods model (LS-COM). The LS-COM modeling frame incorporates the advantages of SEM, multilevel modeling, longitudinal modeling, and MTMM modeling with the flexibility of interchangeable and structurally different methods (Coulacoglou & Saklofske, 2017). For details refer also to Koch et al., 2018.

5. Summary and Conclusion

MTMM was developed in an attempt to assess: 1) the relationship between the same construct (trait) and the same methods of measurement (with the reliabilities in the diagonal in **Table 3**); 2) the relationship between the same construct (trait) using different measurement (i.e., convergent validity); and 3) the relationship between different constructs using different methods of measurement (i.e., discriminant validity), as commented by Price (2017).

The classic MTMM theory encompasses the following concepts. The mono-trait-monomethod blocks comprise the correlations between the traits, measured by one common method. The mono-trait-heteromethod block contains the correlations between the traits measured by different methods, to examine convergent validity. The higher the correlations in the mono-trait-heteromethod block the higher the convergent validity. The heterotrait-monomethod block refers to correlations between different traits measured by the one common method to evaluate the discriminant validity of the measures. The higher the correlations in the heterotrait-monomethod block the lowest the discriminant validity. Finally, the heterotrait-heteromethod block includes correlations of different traits assessed by different methods, also indicating discriminant validity (Koch et al., 2018). These four data arrangements, show different correlations levels, that can be ranked from the highest to the lowest as follows: 1) the reliability coefficients in the diagonals of the monomethod blocks; 2) the validity coefficients in the diagonals of the heteromethod blocks; 3) the heterotrait monomethod coefficients; and 4) heterotrait heteromethod coefficients; Note that 3) and 4) are off-diagonal elements (Sawilowsky, 2007). See for example the multitrait-multimethod matrix in **Table 2**.

CFA methods to MTMM matrices are to some extent underutilized in the empirical literature, and the same is true for the “classic” MTMM approach in general. Probably, this is to some extent due to the lack of availability of multiple measures for a given construct (Brown, 2015). For example, as a minimum, three assessment methods must be typically available for the CU method. Additionally, MTMM methods can be also employed for constructs assessed by a single assessment method, e.g. parenting dimensions measured by different self-report questionnaires (c.f. Green, Goldman and Salovey, 1993 as quoted by Brown, 2015). CFA MTMM models can be easily used in tandem with other methods. As Brown (2015), comments they can be integrated into structural equation models or MIMIC models. In this way, trait factors could be related to background variables in the role of external validators establishing, for example, the predictive validity of traits. A second possible extension is to integrate an MTMM model across different groups to test whether trait factors show measurement invariance (Brown, 2015).

As originally theorized, evidence of convergence and discriminant validity can be shown by correlations between scores of various instruments. In Campbell and Fiske’s own words, (1959: p. 104), “*Convergence between independent measures of the same trait and discrimination between measures of different traits*”. In a similar vein, the current version of the APA standards (APA/AERA, 1999, 2014) indicates that discriminant and convergence validity evidence should be established through correlations of test scores with scores from different instruments and not with different subscales, as noted by Gunnell et al. (2014). Almost 60 years after, the CFA framework renders the MTMM method an equally valuable tool to the classic MTMM when first introduced (Koch et al., 2018).

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for Educational and Psychological Testing* (2nd ed.). Washington, DC: Authors.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing* (3rd ed.). Washington, DC: Authors.
- Bagozzi, R. P. (1993). Assessing Construct Validity in Personality Research: Applications to Measures of Self-Esteem. *Journal of Research in Personality, 27*, 49-87.
- Bagozzi, R. P., & Yi, Y. (1990). Assessing METHOD variance in Multitrait-Multimethod Matrices: The Case of Self-Reported Affect and Perceptions at Work. *Journal of Applied Psychology, 75*, 547-560.
- Benson, J. (1988). Developing a Strong Program of Construct Validation: A Test Anxiety Example. *Educational Measurement: Issues and Practice, 17*, 10-17.

- Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (2nd ed.). New York: The Guilford Press.
- Brown, T. A., & Moore, M. T. (2012). Confirmatory Factor Analysis. In R. H. Hoyle (Ed.), *Handbook of Structural Equation Modeling* (pp. 361-379). New York, NY: Guilford Publications.
- Browne, M. W. (1984). The Decomposition of Multitrait-Multimethod Matrices. *British Journal of Mathematical and Statistical Psychology*, *37*, 1-21.
- Byrne, B. M. (2012). *Structural Equation Modeling with Mplus: Basic Concepts, Applications, and Programming* (2nd ed.). New York: Routledge.
- Byrne, B. M., & Bazana, P. G. (1996). Investigating the Measurement of Social and Academic Competencies for Early/Late Preadolescents and Adolescents: A Multitrait-Multimethod Analysis. *Applied Measurement in Education*, *9*, 113-132.
- Byrne, B. M., & Goffin, R. D. (1993). Modeling MTMM Data from Additive and Multiplicative Covariance Structures: An Audit of Construct Validity Concordance. *Multivariate Behavioral Research*, *28*, 67-96.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix. *Psychological Bulletin*, *56*, 81-105.
- Campbell, D. T., & O'Connell, E. J. (1967). Method Factors in Multitrait-Multimethod Matrices: Multiplicative Rather than Additive? *Multivariate Behavioral Research*, *2*, 409-426.
- Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling*, *14*, 464-504.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling*, *9*, 233-255.
- Coenders, G., & Saris, W. E. (2000). Testing Nested Additive, Multiplicative, and General Multitrait-Multimethod Models. *Structural Equation Modeling*, *7*, 219-250. https://doi.org/10.1207/S15328007SEM0702_5
- Cole, D. A. (1987). Utility of Confirmatory Factor Analysis in Test Validation Research. *Journal of Consulting and Clinical Psychology*, *55*, 584-594. <https://doi.org/10.1037/0022-006X.55.4.584>
- Cole, D. A., Ciesla, J. A., & Steiger, J. (2007). The Insidious Effects of Failing to Include Design-Driven Correlated Residuals in Latent-Variable Covariance Structure Analysis. *Psychological Methods*, *12*, 381-398. <https://doi.org/10.1037/1082-989X.12.4.381>
- Cole, D. A., Howard, G. S., & Maxwell, S. E. (1981). Effects of Mono- versus Multiple-Operationalization in Construct Validation Efforts. *Journal of Consulting and Clinical Psychology*, *49*, 395-405. <https://doi.org/10.1037/0022-006X.49.3.395>
- Conway, J. M., Scullen, S. E., Lievens, F., & Lance, C. E. (2004). Bias in the Correlated Uniqueness Model for MTMM Data. *Structural Equation Modeling*, *11*, 535-559. https://doi.org/10.1207/s15328007sem1104_3
- Coover, M. D., Craiger, J. P., & Teachout, M. S. (1997). Effectiveness of the Direct Product versus Confirmatory Factor Model for Reflecting the Structure of Multimethod-Multirater Job Performance Data. *Journal of Applied Psychology*, *82*, 271-280. <https://doi.org/10.1037/0021-9010.82.2.271>
- Corten, I. W., Saris, W. E., Coenders, G., van der Veld, W., Aalberts, C. E., & Kornelis, C. (2002). Fit of Different Models for Multitrait-Multimethod Experiments. *Structural Equation Modeling*, *9*, 213-232. https://doi.org/10.1207/S15328007SEM0902_4
- Coulacoglou, C., & Saklofske, D. H. (2017). *Psychometrics and Psychological Assessment: Principles and Applications*. Oxford, UK: Academic Press.

- Courvoisier, D. S., Nussbeck, F. W., Eid, M., Geiser, C., & Cole, D. A. (2008). Analyzing the Convergent and Discriminant Validity of States and Traits: Development and Applications of Multimethod Latent State-Trait Models. *Psychological Assessment, 20*, 270-280. <https://doi.org/10.1037/a0012812>
- Cronbach, L. J., & Meehl, P. E. (1955). Construct Validity in Psychological Tests. *Psychological Bulletin, 52*, 281-302. <https://doi.org/10.1037/h0040957>
- Cudeck, R. (1988). Multiplicative Models and MTMM Matrices. *Journal of Educational Statistics, 13*, 131-147. <https://doi.org/10.2307/1164750>
- DeVellis, R. F. (2017). *Scale Development: Theory and Applications* (4th ed.). Thousand Oaks, CA: Sage.
- Dudgeon, P. (1994). A Reparameterization of the Restricted Factor Analysis Model for Multitrait-Multimethod Matrices. *British Journal of Mathematical and Statistical Psychology, 47*, 283-308. <https://doi.org/10.1111/j.2044-8317.1994.tb01038.x>
- Dumenci, L. (2000). Multitrait-Multimethod Analysis. In S. D. Brown, & H. E. A. Tinsley (Eds.), *Handbook of Applied Multivariate Statistics and Mathematical Modeling* (pp. 583-611). San Diego, CA: Academic Press. <https://doi.org/10.1016/B978-012691360-6/50021-5>
- Eid, M. (2000). A Multitrait-Multimethod Model with Minimal Assumptions. *Psychometrika, 65*, 241-261. <https://doi.org/10.1007/BF02294377>
- Eid, M. (2010). Multitrait-Multimethod-Matrix. In N. Salkind (Ed.), *Encyclopedia of Research Design* (Vol. 1, pp. 850-855). Thousand Oaks, CA: Sage.
- Eid, M., Lischetzke, T., & Nussbeck, F. W. (2006). Structural Equation Models for Multitrait-Multimethod Data. In M. Eid, & E. Diener (Eds.), *Handbook of Multimethod Measurement in Psychology* (pp. 283-299). Washington DC: American Psychological Association. <https://doi.org/10.1037/11383-020>
- Eid, M., Lischetzke, T., Nussbeck, F. W., & Trierweiler, L. I. (2003). Separating Trait Effects from Trait-Specific Method Effects in Multitrait-Multimethod Models: A Multiple-Indicator CT-C (M-1) Model. *Psychological Methods, 8*, 38-60. <https://doi.org/10.1037/1082-989X.8.1.38>
- Eid, M., Nussbeck, F. W., Geiser, C., Cole, D. A., Gollwitzer, M., & Lischetzke, T. (2008). Structural Equation Modeling of Multitrait-Multimethod Data: Different Models for Different Types of Methods. *Psychological Methods, 13*, 230-253. <https://doi.org/10.1037/a0013219>
- Flamer, S. (1983). Assessment of the Multitrait-Multimethod Matrix Validity of Likert Scales via Confirmatory Factor Analysis. *Multivariate Behavioral Research, 18*, 275-308. https://doi.org/10.1207/s15327906mbr1803_3
- Franzen, M. D. (2002). *Reliability and Validity in Neuropsychological Assessment* (2nd ed.). New York: Springer. <https://doi.org/10.1007/978-1-4757-3224-5>
- Furr, R. M. (2011). *Scale Construction and Psychometrics for Social and Personality Psychology*. New Delhi, IN: Sage Publications. <https://doi.org/10.4135/9781446287866>
- Geiser, C., Eid, M., & Nussbeck, F. W. (2008). On the Meaning of the Latent Variables in the CTC (M-1) Model: A Comment on Maydeu-Olivares and Coffman (2006). *Psychological Methods, 13*, 49-57. <https://doi.org/10.1037/1082-989X.13.1.49>
- Geiser, C., Eid, M., West, S. G., Lischetzke, T., & Nussbeck, F. W. (2012). A Comparison of Method Effects in Two Confirmatory Factor Models for Structurally Different Methods. *Structural Equation Modeling, 19*, 409-436. <https://doi.org/10.1080/10705511.2012.687658>
- Goffin, R. D., & Jackson, D. N. (1992). Analysis of Multitrait-Multirater Performance

- Appraisal Data: Composite Direct Product Method versus Confirmatory Factor Analysis. *Multivariate Behavioral Research*, 27, 363-385.
https://doi.org/10.1207/s15327906mbr2703_4
- Green, D. P., Goldman, S. L., & Salovey, P. (1993). Measurement Error Masks Bipolarity in Affect Ratings. *Journal of Personality and Social Psychology*, 64, 1029-1041.
<https://doi.org/10.1037/0022-3514.64.6.1029>
- Grimm, K. J., Pianta, R. C., & Konold, T. (2009). Longitudinal Multitrait-Multimethod Models for Developmental Research. *Multivariate Behavioural Research*, 44, 233-258.
<https://doi.org/10.1080/00273170902794230>
- Gunnell, K. E., Schellenberg, B. J. I., Wilson, P. M., Crocker, P. R. E., Mack, D. E., & Zumbo, B. D. (2014). A Review of Validity Evidence Presented in the Journal of Sport and Exercise Psychology (2002-2012): Misconceptions and Recommendations for Validation Research. In B. D. Zumbo, & E. K. H. Chan (Eds.), *Validity and Validation in Social, Behavioral, and Health Sciences, Social Indicators Research Series 54* (pp. 137-156). Switzerland: Springer International.
https://doi.org/10.1007/978-3-319-07794-9_8
- Hernández, A., & González-Romá, V. (2002). Analysis of Multitrait-Multioccasion Data: Additive versus Multiplicative Models. *Multivariate Behavioral Research*, 37, 59-87.
https://doi.org/10.1207/S15327906MBR3701_03
- Hox, J. J., & Kleiboeer, A. M. (2007). Retrospective Questions or a Diary Method? A Two-Level Multitrait-Multimethod Analysis. *Structural Equation Modeling*, 14, 311-325. <https://doi.org/10.1080/10705510709336748>
- Hu, L. T., & Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives. *Structural Equation Modeling*, 6, 1-55. <https://doi.org/10.1080/10705519909540118>
- Jackson, D. N. (1969). Multimethod Factor Analysis in the Evaluation of Convergent and Discriminant Validity. *Psychological Bulletin*, 72, 30-49.
<https://doi.org/10.1037/h0027421>
- Kenny, D. A. (1976). An Empirical Application of Confirmatory Factor Analysis to the Multitrait-Multimethod Matrix. *Journal of Experimental Social Psychology*, 12, 247-252. [https://doi.org/10.1016/0022-1031\(76\)90055-X](https://doi.org/10.1016/0022-1031(76)90055-X)
- Kenny, D. A. (1979). *Correlation and Causality*. New York: Wiley.
- Kenny, D. A., & Kashy, D. A. (1992). Analysis of the Multitrait-Multimethod Matrix by Confirmatory Factor Analysis. *Psychological Bulletin*, 112, 165-172.
<https://doi.org/10.1037/0033-2909.112.1.165>
- Kline, R. B. (2009). *Becoming a Behavioral Science Researcher: A Guide to Producing Research That Matters*. New York: Guilford Publications.
- Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (4th ed.). New York: Guilford Publications.
- Koch, T., Eid, M., & Lochner, K. (2018). Multitrait-Multimethod-Analysis: The Psychometric Foundation of CFA-MTMM Models. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development* (VII, pp. 781-846). Hoboken, NJ: Wiley.
- Koch, T., Schultze, M., Eid, M., & Geiser, C. (2014). A Longitudinal Multilevel CFA-MTMM Model for Interchangeable and Structurally Different Methods. *Frontiers in Psychology*, 5, 311. <https://doi.org/10.3389/fpsyg.2014.00311>
- Kyriazos, T. A. (2018). Applied Psychometrics: The 3-Faced Construct Validation Method, a Routine for Evaluating a Factor Structure. *Psychology*, 9, 2044-2072.

- <https://doi.org/10.4236/psych.2018.98117>
- Kyriazos, T. A., Stalikas, A., Prassa, K., & Yotsidi, V. (2018a). Can the Depression Anxiety Stress Scales Short Be Shorter? Factor Structure and Measurement Invariance of DASS-21 and DASS-9 in a Greek, Non-Clinical Sample. *Psychology, 9*, 1095-1127. <https://doi.org/10.4236/psych.2018.95069>
- Kyriazos, T. A., Stalikas, A., Prassa, K., & Yotsidi, V. (2018b). A 3-Faced Construct Validation and a Bifactor Subjective Well-Being Model Using the Scale of Positive and Negative Experience, Greek Version. *Psychology, 9*, 1143-1175. <https://doi.org/10.4236/psych.2018.95071>
- Kyriazos, T. A., Stalikas, A., Prassa, K., Galanakis, M., Flora, K., & Chatzilia, V. (2018c). The Flow Short Scale (FSS) Dimensionality and What MIMIC Shows on Heterogeneity and Invariance. *Psychology, 9*, 1357-1382. <https://doi.org/10.4236/psych.2018.96083>
- Kyriazos, T. A., Stalikas, A., Prassa, K., Galanakis, M., Yotsidi, V., & Lakioti, L. (2018d). Psychometric Evidence of the Brief Resilience Scale (BRS) and Modeling Distinctiveness of Resilience from Depression and Stress. *Psychology, 9*, 1828-1857.
- Kyriazos, T. A., Stalikas, A., Prassa, K., Yotsidi, V., Galanakis, M., & Pezirkianidis, C. (2018e). Validation of the Flourishing Scale (FS), Greek Version and Evaluation of Two Well-Being Models. *Psychology, 9*, 1789-1813. <https://doi.org/10.4236/psych.2018.97105>
- LaGrange, B., & Cole, D. A. (2008). An Expansion of the Trait-State-Occasion Model: Accounting for Shared Method Variance. *Structural Equation Modeling, 15*, 241-271. <https://doi.org/10.1080/10705510801922381>
- Lance, C. E., Noble, C. L., & Scullen, S. E. (2002). A Critique of the Correlated Trait Correlated Method and Correlated Uniqueness Models for Multitrait-Multimethod Data. *Psychological Methods, 7*, 228-244. <https://doi.org/10.1037/1082-989X.7.2.228>
- Lievens, F., & Conway, J. M. (2001). Dimension and Exercise Variance in Assessment Center Scores: A Large-Scale Evaluation of Multitrait-Multimethod Matrices. *Journal of Applied Psychology, 86*, 1202-1222. <https://doi.org/10.1037/0021-9010.86.6.1202>
- Loehlin, J. C. (2004). *Latent Variable Models* (4th ed.). Mahwah, NJ: Erlbaum.
- Loehlin, J. C., & Beaujean, A. A. (2017). *Latent Variable Models: An Introduction to Factor, Path, and Structural Equation Analysis*. New York: Taylor & Francis.
- Marsh, H. W. (1988). Multitrait-Multimethod Analyses. In J. P. Keeves (Ed.), *Educational Research Methodology, Measurement, and Evaluation: An International Handbook* (pp. 570-578). Oxford: Pergamon.
- Marsh, H. W. (1989). Confirmatory Factor Analysis of Multitrait-Multimethod Data: Many Problems and a Few Solutions. *Applied Psychological Measurement, 13*, 335-361. <https://doi.org/10.1177/014662168901300402>
- Marsh, H. W., & Bailey, M. (1991). Confirmatory Factor Analyses of Multitrait-Multimethod Data: A Comparison of Alternative Models. *Applied Psychological Measurement, 15*, 47-70. <https://doi.org/10.1177/014662169101500106>
- Marsh, H. W., & Byrne, B. M. (1993). Confirmatory Factor Analysis of Multitrait-Multimethod Self-Concept Data: Between-Group and Within-Group Invariance Constraints. *Multivariate Behavioral Research, 28*, 313-349. https://doi.org/10.1207/s15327906mbr2803_2
- Marsh, H. W., & Grayson, D. (1995). Latent Variable Models of Multitrait-Multimethod Data. In R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, Issues, and Applications* (pp. 177-198). Thousand Oaks, CA: Sage.
- Marsh, H. W., & Hocevar, D. (1983). Confirmatory Factor Analysis of Multi-

- trait-Multimethod Matrices. *Journal of Educational Measurement*, *20*, 231-248.
<https://doi.org/10.1111/j.1745-3984.1983.tb00202.x>
- Marsh, H. W., Byrne, B. M., & Craven, R. (1992). Overcoming Problems in Confirmatory Factor Analyses of MTMM Data: The Correlated Uniqueness Model and Factorial Invariance. *Multivariate Behavioral Research*, *27*, 489-507.
https://doi.org/10.1207/s15327906mbr2704_1
- Martin, A. J. & Marsh, H. W. (2006). Academic Resilience and Its Psychological and Educational Correlates: A Construct Validity Approach. *Psychology in the Schools*, *43*, 267-281. <https://doi.org/10.1002/pits.20149>
- Maydeu-Olivares, A., & Coffman, D. L. (2006). Random Intercept Item Factor Analysis. *Psychological Methods*, *11*, 344-362. <https://doi.org/10.1037/1082-989X.11.4.344>
- Messick, S. (1980). Test Validity and the Ethics of Assessment. *American Psychologist*, *35*, 1012-1027. <https://doi.org/10.1037/0003-066X.35.11.1012>
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed., (pp. 13-103). New York: Macmillan Publishing Co.
- Messick, S. (1995). Validity of Psychological Assessment: Validation of Inferences from Persons' Responses and Performances as Scientific Inquiry into Score Meaning. *American Psychologist*, *50*, 741-749. <https://doi.org/10.1037/0003-066X.50.9.741>
- Meyer, J. F., Frost, R. O., Brown, T. A., Steketee, G., & Tolin, D. F. (2013). A Multi-trait-Multimethod Matrix Investigation of Hoarding. *Journal of Obsessive-Compulsive and Related Disorders*, *2*, 273-280. <https://doi.org/10.1016/j.jocrd.2013.03.002>
- Millsap, R. E. (1995a). Measurement Invariance, Predictive Invariance, and the Duality Paradox. *Multivariate Behavioral Research*, *30*, 577-605.
https://doi.org/10.1207/s15327906mbr3004_6
- Millsap, R. E. (1995b). The Statistical Analysis of Method Effects in Multi-trait-Multimethod Data: A Review. In P. E. Shrout, & S. T. Fiske (Eds.), *Personality Research, Methods, and Theory: A Festschrift Honoring Donald W. Fiske* (pp. 93-109). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pohl, S., & Steyer, R. (2010). Modeling Common Traits and Method Effects in Multi-trait-Multimethod Analysis. *Multivariate Behavioral Research*, *45*, 45-72.
<https://doi.org/10.1080/00273170903504729>
- Pohl, S., Steyer, R., & Kraus, K. (2008). Modelling Method Effects as Individual Causal Effects. *Journal of the Royal Statistical Society Series A*, *171*, 41-63.
- Price, L. R. (2017). *Psychometric Methods: Theory into Practice*. New York: The Guilford Press.
- Revelle, W. (2018). *Using the Psych Package to Generate and Test Structural Models*.
<https://personality-project.org/r/psych-manual.pdf>
- Saris, W. E., & Aalberts, C. (2003). Different Explanations for Correlated Disturbance Terms in MTMM Studies. *Structural Equation Modeling*, *10*, 193-213.
https://doi.org/10.1207/S15328007SEM1002_2
- Sawilowsky, S. (2002). A Quick Distribution-Free Test Fortrend That Contributes Evidence of Construct Validity. *Measurement and Evaluation in Counseling and Development*, *35*, 78-88.
- Sawilowsky, S. S. (2007). Construct Validity. In N. J. Salkind (Ed.), *Encyclopedia of Measurement and Statistics* (pp. 178-180). Thousand Oaks, CA: Sage.
- Schmitt, N., & Stults, D. M. (1986). Methodology Review: Analysis of Multi-trait-Multimethod Matrices. *Applied Psychological Measurement*, *10*, 1-22.

- <https://doi.org/10.1177/014662168601000101>
- Shrout, P. E., & Fiske, S. T. (Eds.) (1995). *Personality Research, Methods, and Theory: A Festschrift Honoring Donald W. Fiske*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Steyer, R. (1989). Models of Classical Psychometric Test Theory as Stochastic Measurement Models: Representation, Uniqueness, Meaningfulness, Identifiability, and Testability. *Methodika*, 3, 25-60.
- Steyer, R. (2005). Analyzing Individual and Average Causal Effects via Structural Equation Models. *Methodology*, 1, 39-54. <https://doi.org/10.1027/1614-1881.1.1.39>
- Steyer, R., Partchev, I., & Shanahan, M. J. (2000). Modeling True Intraindividual Change in Structural Equation Models: The Case of Poverty and Children's Psychosocial Adjustment. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling Longitudinal and Multilevel Data: Practical Issues, Applied Approaches, and Specific Examples* (pp. 109-126). Mahwah, NJ: Lawrence Erlbaum.
- Tomás, J. M., Hontangas, P. M., & Oliver, A. (2000). Linear Confirmatory Factor Models to Evaluate Multitrait-Multimethod Matrices: The Effects of Number of Indicators and Correlation among Methods. *Multivariate Behavioral Research*, 35, 469-499. https://doi.org/10.1207/S15327906MBR3504_03
- Trochim, W. M. (2002). *Research Methods Knowledge Base*. Ithaca, NY: Cornell University.
- Verhees, J., & Wansbeek, T. J. (1990). A Multimode Direct Product Model for Covariance Structure Analysis. *British Journal of Mathematical and Statistical Psychology*, 43, 231-240. <https://doi.org/10.1111/j.2044-8317.1990.tb00938.x>
- Widaman, K. F. (1985). Hierarchically Nested Covariance Structure Models for Multitrait-Multimethod Data. *Applied Psychological Measurement*, 9, 1-26. <https://doi.org/10.1177/014662168500900101>
- Wothke W., & Browne, M. W. (1990). The Direct Product Model for the MTMM Matrix Parameterized as a Second Order Factor Analysis Model. *Psychometrika*, 55, 255-262. <https://doi.org/10.1007/BF02295286>
- Wothke, W. (1993). Nonpositive Definite Matrices in Structural Modeling. In K. A. Bollen, & J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 256-293). Newbury Park, CA: Sage.
- Wothke, W. (1995). Covariance Components Analysis of the Multitrait-Multimethod Matrix. In P. E. E. A. Shrout (Ed.), *Personality Research, Methods, and Theory: A Festschrift Honoring Donald W. Fiske* (pp. 125-144). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Zimmerman, D. W. (1975). Two Concepts of "True-Score" in Test Theory. *Psychological Reports*, 36, 795-805.
- Zinbarg, R. E., Yovel, I., Revelle, W., & McDonald, R. P. (2006). Estimating Generalizability to a Latent Variable Common to All of a Scale's Indicators: A Comparison of Estimators for ω_h . *Applied Psychological Measurement*, 30, 121-144. <https://doi.org/10.1177/0146621605278814>