Browne's Composite Direct Product Model for MTMM Correlation Matrices: A Review

Timo M Bechger

Abstract.

This paper concerns the analysis of the multitrait-multimethod correlation matrix with the composite direct product model proposed by Browne (1984). The paper contains a review of the literature related to the composite direct product model and provides guidelines to interpretation and analysis.
1. Introduction

According to Campbell and Fiske (1959), there are two aspects to construct validity: one is *convergent validation*, which requires agreement between scores obtained with two or more different measures of the same trait or construct. The other aspect requires a demonstration that the construct can be differentiated from other traits, and to demonstrate this a researcher must show disagreement between two scores that presumably measure different constructs. Such disagreement is evidence of *discriminant validity*. When there are two tests measuring the same trait this means that tests can be shown to be invalid, not only because of low correlations but also because of high correlations with other tests purporting to measure different things (Campbell & Fiske, 1959, p. 84).

To facilitate the comparison of different kinds of correlations, Campbell and Fiske (1959) arranged the correlations in a *MultiTrait-MultiMethod (MTMM) matrix*. The MTMM-matrix is a correlation matrix between measurements under combinations of traits T and methods M. If there are t traits and m methods, $p = t \times m$ measurements are made on each of N subjects.

Campbell and Fiske (1959) suggested that four criteria should hold for measures that have convergent and/or discriminant validity:

1. The monotrait-heteromethod correlations or *validity values* are large.
2. The validity values should exceed the absolute values of correlations in the row and column of their respective heteromethod submatrices.
3. The validity values must be higher than the correlations in each of the two corresponding monomethod submatrices.
4. All submatrices of intertrait correlations should have the same pattern, no matter which methods are used.
A number of formal models have been proposed to (a) describe MTMM matrices in the presence of measurement error and (b) to quantify and extend the Campbell and Fiske conditions (Schmitt and Stulz, 1986; Wothke, 1995). Currently, it is common practice to impose upon the MTMM matrix an additive covariance structure model using confirmatory or restricted factor analysis (Jöreskog, 1973; Widaman, 1985).

Campbell and O'Connell (1967, 1982) came to the conclusion that the method effects were multiplicative rather than additive. Swain (1975) suggested a direct product model that is suitable for the analysis of MTMM covariance matrices that have the multiplicative property observed by Campbell and O'Connell. (See also Verhees and Wansbeek, 1990). The Composite Direct Product (CDP) model (Browne, 1984) is an extension of Swain's model which is suitable for correlation matrices and allows measurement error in the observations.

The purpose of this paper is twofold: First, we present a review of the literature related to the CDP model. Second, we wish to discuss the interpretation and the practical application of the CDP model in more detail then is usually done.

2. The Work of Campbell and O'Connell
Monomethod correlations are normally higher than heteromethod correlations. Usually it is assumed that sharing the same method augments inter trait relationships above the true values, which are more validly seen in heteromethod form. Assuming that this is true Campbell and O'Connell (1967, 1982) investigated the relation between heteromethod and monomethod correlations. They plotted monomethod correlations against heteromethod correlations and calculated the slope of the linear regression in each of the plots. If shared method elements add to the correlations in a linear manner, they argue, the heteromethod correlations are an additive function of the
monomethod correlations and the relation will best be described by a straight
line with a slope that is smaller or equal to 1.00 which equals the 45° line if
there is no method effect. What they found however was that the relationship
was best described by a straight line, passing through the origin with a slope
greater than 1.00. Campbell and O'Connell concluded that the argument that
method factors augment the intertrait correlations is correct but the
magnitude depends on the level of true trait relationship. The higher the true
trait relationship, the more this relationship is augmented.

3. The Composite Direct Product Model for MTMM Matrices

Let \( x(T_j M_k) \) denote the random variable obtained by combining the j-th trait
with the k-th method.\(^1\)

\[
 x(T_j, M_k) = \mu(T_j, M_k) + \delta(T_j, M_k) \times \{ c(T_j, M_k) + U(T_j, M_k) \} \quad (1)
\]

where:

\( \mu(T_j, M_k) \) is the mean of the distribution of \( x(T_j, M_k) \),
\( \delta(T_j, M_k) \) is a scaling factor,
\( c(T_j, M_k) \) is a zero-mean random variate representing the common
part of \( x(T_j, M_k) \), and
\( U(T_j, M_k) \) is a zero-mean random variate representing the unique (or
error) part of \( x(T_j, M_k) \).

If we assume that the common part is uncorrelated to the unique term,
the expected mean and covariance of the observed measures are

\(^1\) The notation is chosen to be compatible with Browne's notation.
\begin{align*}
E \left[ x(T_j, M_k) \times x(T_g, M_s) \right] &= \delta(T_j, M_k) \times \left\{ E \left[ c(T_j, M_k) \times c(T_g, M_s) \right] \right\} + \\
&+ E \left[ U(T_j, M_k) \times U(T_g, M_s) \right] \times \delta(T_g, M_s) 
\end{align*}

(2)

and

\begin{equation}
E \left[ x(T_j, M_k) \right] = \mu(T_j, M_k)
\end{equation}

(3)

In matrix notation Equation (2) may be written as:

\begin{equation}
\Sigma = \Delta (\Sigma_c + \Sigma_u) \Delta
\end{equation}

(4)

where $\Delta = \{ \delta(T_j, M_k) \}$ is a $p \times p$ diagonal scaling matrix, $\Sigma_c$ is the $p \times p$ common score covariance matrix and $\Sigma_u$ is the $p \times p$ unique score covariance matrix. $\Sigma$ is the (implied) correlation or covariance matrix in the population.

The CDP model may be derived from the following multiplicative structure for $c(T_j, M_k)$:

\begin{equation}
c(T_j, M_k)_i = z(T_j)_i \times z(M_k)_i
\end{equation}

(5)

It is assumed that the response of subject $i$ to the item in the multitrait-multimethod design is the product of two latent variables.
In matrix notation the common score model in (5) becomes

\[ c = z(M) \otimes z(T) = \]

where \( c = [ c(T_1, M_1), c(T_2, M_1), \ldots, c(T_t, M_m) ] \), \( z(T)' = [ z(T_1), z(T_2), \ldots, z(T_t) ] \), \( z(M)' = [ z(M_1), z(M_2), \ldots, z(M_m) ] \) and \( \otimes \) symbolizes the (right) Kronecker or direct product operator. The individual true scores are regarded as independent realizations of the random vector variate \( c \) and the variables are ordered such that traits are nested within methods.

The matrix notation of the common score covariance matrix can be found by expanding the expected value of \( cc' \), denoted by \( E[cc'] \):

\[
E[ ( z(M) \otimes z(T) ) ( z(M) \otimes z(T) )' ] = \\
E[ z(M) \otimes z(T) ( z(M)' \otimes z(T)' ) ] (7)
\]

The expected value involves fourth-order moments in terms of the \( z(T) \) and \( z(M) \) variables. The fourth-order cumulant is zero if we assume that \( z(T) \) and \( z(M) \) are normally distributed. If, in addition, we assume that \( z(T) \) and \( z(M) \) are independent, Equation (7) can be simplified and written as:
\[ E[z(M)z(M)'] \otimes E[z(T)z(T)'] \]
\[ \Phi_M \otimes \Phi_T \]

(8)

The common dispersion matrix is standardized by pre- and post-multiplying the dispersion matrices by a diagonal matrix which contains the inverse of the common score standard deviations. From Equation (8) this matrix is found to be

\[ D = \text{diag}^{-1} (\Sigma_c) = D_M \otimes D_T \]

(9)

where \( D_T = \text{diag}^{-1} (\Phi_T) \) and \( D_M = \text{diag}^{-1} (\Phi_M) \). With Eqs. (8) and (9) we now derive the common score correlation matrix \( P_c \):

\[ D \Sigma_c D = (D_M \Phi_M D_M) \otimes (D_T \Phi_T D_T) = P_c = P_M \otimes P_T \]

(10)

The symmetric matrices \( P_T \) (\( t \times t \)), \( P_M \) (\( m \times m \)) are the trait and method correlation matrices. These matrices have the mathematical properties of correlation matrices whose elements represent multiplicative components of common factor score correlations or, 'correlations corrected for attenuation'. With two trait and two methods the expanded form of (10) is equal to

\[ P_c = \left[ \begin{array}{ccc} 1 & \rho_c(T_1,T_2) & 1 \\ \rho_c(T_1,T_2) & 1 & \rho_c(T_2,T_1) \\ \rho_c(M_2,M_1) & \rho_c(T_2,T_1) & \rho_c(T_1,T_2) \end{array} \right] \]

(11)
Inspection of Equation (11) shows that the CDP model assumes that the correlation matrix corrected for attenuation has the following properties:

1. The diagonal submatrices are all equal to $P_T$.
2. The validity values are all equal to the associated elements of $P_M$.
3. The heteromethod heterotrait correlations are the product of terms $\rho_c(M_r, M_s)$ and $\rho_c(T_g, T_j)$.
4. The heterotrait correlations are symmetric around the diagonal.
5. The CDP model possess the specific multiplicative property that was observed by Campbell and O'Connell. The relation between heterotrait heteromethod values and heterotrait monomethod correlations is given by:

$$\rho_c(T_g M_r, T_j M_r) = \rho_c(T_g M_r, T_j M_s) \times \{ \rho_c(M_r, M_s) \}^{-1}$$  \hspace{1cm} (12)

Since $\rho_c(M_r, M_s) \leq 1.00$ the slope $\{ \rho_c(M_r, M_s) \}^{-1}$ is equal or bigger than $|1.00|$, depending on the sign of $\rho_c(M_r, M_s)$ (Browne, 1984, Equation 4.7).

6. Monomethod correlations equal $\rho_c(T_g, T_j)$ and heteromethod correlations are $\rho_c(T_g, T_j) \times \rho_c(M_s, M_r)$ which implies that the off-diagonal entries in $P_M$ indicate changes in MTMM matrix patterns relative to the monomethod block. For example, a value of 0.72 in $P_M$ signifies a 28 percent drop in the correlation coefficient when the traits are measured with different methods. Similarly, off-diagonal entries in $P_T$ express magnitude changes relative to validity values. If $\hat{\rho}_c(T_2, T_1) = 0.69$ the average correlation between $T_1$ and $T_2$ is 31 percent smaller than the convergent validity values (cf. Wothke, 1995, p. 24).
The multiplicative data model in (4) and Equation (10) generate the model for the observed correlation matrix:

$$\Delta \Sigma \Delta = \Delta (D^{-1} P_c D^{-1} + \Sigma_u) \Delta = \Delta^* (P_M \otimes P_T + \Sigma_u D^2) \Delta^*$$  \hspace{1cm} (13)

Where $\Delta^* = \Delta D^{-1}$, with $D$ as in Equation (9), and $\Delta$ is a diagonal matrix that contains the reciprocals of the sample standard deviations. To ensure the scale independence of the elements of $\Sigma_u$ the model is frequently reparameterized as follows (Cudeck, 1988, p. 139):

$$\Delta^* (P_M \otimes P_T + U^2) \Delta^*$$  \hspace{1cm} (14)

where $\Delta^* = \Delta D^{-1}$, $D$ as in Equation (9), and the $p \times p$ diagonal matrix $U^2$ represents the ratios of unique variance to common score variance i.e.,

$$U^2 = \Sigma_u D^2 = \left\{ \frac{\text{var}(U) \times 1}{\text{var}(c)} \right\}.$$  \hspace{1cm} (15)

The CDP model is identified with two traits and two methods or more. This can be proved by equating Equation (11) with the observed correlation matrix (Cudeck 1988, p. 138). The number of degrees of freedom is

$$df = \left\{ p(p-3) - t(t-1) - m(m-1) \right\} / 2$$  \hspace{1cm} (16)

Equation (14) shows that the model is scale free in the sense that scale differences among the observed measures are absorbed by the scaling terms in the matrix $\Delta^*$ and the model is not affected by differences in scale. This is an important property of the model since in the social sciences scale information
is typically regarded as arbitrary or of little interest and it is customary to analyse the correlation matrix.

4. The CDP Model and the Campbell and Fiske Criteria

When the CDP model fits the data its parameter values provide information as to whether or not correlation coefficients corrected for attenuation meet the four conditions of Campbell and Fiske (Browne, 1984 pp. 9-10; Cudeck, 1988, pp. 136-137: see Table 1).

Table 1: The relation of parameters in the CDP model with the four criteria of Campbell and Fiske (1959).

<table>
<thead>
<tr>
<th>Campbell and Fiske Criteria</th>
<th>Conditions to be met in the CDP model.</th>
</tr>
</thead>
<tbody>
<tr>
<td>First of all the CDP model must be accepted.</td>
<td>Method correlations are substantial.</td>
</tr>
<tr>
<td>Criterion 1</td>
<td>All trait correlations are small and do not approach plus or minus unity.</td>
</tr>
<tr>
<td>Criterion 2</td>
<td>The method correlations are bigger than the trait correlations.</td>
</tr>
<tr>
<td>Criterion 3</td>
<td>The CDP model must be accepted.</td>
</tr>
</tbody>
</table>

The first Campbell and Fiske condition requires validity correlations to be large. As validity values equal method correlations this happens when method correlations are substantial. Second, Campbell and Fiske require that the validity values are larger than the off-diagonal elements in the same heteromethod submatrix. The off-diagonal values equal $\rho_c(M_r,M_s) \times \rho_c(T_g,T_j)$, and are lower than the convergent validity values unless $\rho_c(T_g,T_j) = |1.00|$. The third criterion specifies that the validity values should be greater than all monomethod correlations. This is true when elements of $P_M$ are larger than elements of $P_T$. The last criterion is that the pattern of all submatrices of
intertrait correlations is similar. This requirement will obtain whenever the model describes the data adequately (see Equation 11). Under the CDP model:

\[ \rho(T_g M_r, T_g M_r) = \delta^2(T_g M_r) + U^2(T_g M_r) \]

and the scaling terms, \( \delta^2(T_g M_r) \), provide an estimate of the reliability of \( x(T_g M_r) \) (Cudeck, 1988, p. 140).

Finally, note that the elements of \( P_T \) and \( P_M \) can be restricted and the criteria in Table 1 can be tested by comparing the likelihood of the model with and without the restrictions (Bagozzi and Yi, 1992).

5. Parameter Estimates and Goodness-of-Fit

Most programs for structural equation modelling such as LISREL (Jöreskog and Sörbom, 1993) or EQS (Bentler, 1989) have no option for models with Kronecker products. In contrast to these other programs the Mx program recognizes the Kronecker product as a modelling operation so that the CDP model can be specified in its original form (Neale, 1994, 1995). Table 2 presents a general scheme for Mx scripts that is easy to revise for different number of variables.

<table>
<thead>
<tr>
<th>Table 2 : A general Mx setup for the multiplicative model</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE: Browne-Swain CDP model with two facets</td>
</tr>
<tr>
<td>DATA NObservations=n NInput=p NGroup=1</td>
</tr>
<tr>
<td>CM FI=mtmm.cor</td>
</tr>
<tr>
<td>#define methods m</td>
</tr>
<tr>
<td>#define traits t</td>
</tr>
<tr>
<td>#define mxt p</td>
</tr>
<tr>
<td>MATrices</td>
</tr>
<tr>
<td>D DIagonal mxt mxt FREE</td>
</tr>
<tr>
<td>T STandardized traits traits FREE</td>
</tr>
<tr>
<td>M STandardized methods methods FREE</td>
</tr>
<tr>
<td>U DIagonal mxt mxt FREE</td>
</tr>
<tr>
<td>CO D*(M@T + U.U)*D /</td>
</tr>
<tr>
<td>START 1 all</td>
</tr>
<tr>
<td>OU</td>
</tr>
</tbody>
</table>
The italic printed text in Table 2 is optional and the capital letters denote the minimally required input specification.

If one wishes to impose a multiplicative structure on the unique dispersion matrix i.e., $U^2 = U_M \otimes U_T$, this requires only a minor change in the Mx setup. Under these conditions, however, the model is not scale invariant and should not be fitted to the correlation matrix (Browne, 1984, p. 12; Cudeck, 1989). It is equally simple to structure the trait or method correlation matrices (Browne, 1989, p. 16-17) or extend the model to include additional facets (Browne and Strydom, In Press; Cudeck, 1988).

In order to obtain parameter estimates with LISREL8 the CDP model may be written as a common factor model (Dolan and Molenaar, 1992):

$$\Sigma = \Delta P_c \Delta + \Sigma_u = \Lambda \Psi \Lambda' + \Theta_e$$

(17)

Under this parameterization the error variances are estimated directly instead of the ratios of error to the true score variance. $\Theta_e$ (Theta_epsilon) and $\Lambda$ (Lambda) must be specified as diagonal matrices of order $p$. The structure of $\Psi$ (Psi) requires the specification of appropriate non-linear constraints which can be found by writing out the matrix

$$\Psi = P_c = P_M \otimes P_T$$

(18)

as in Equation 11. In the same manner a multiplicative structure can be imposed on $\Theta_e$. (A LISREL8 setup is contained in the appendix). Wothke and Browne (1990; also Wothke, 1995) describe how the CDP model can be fitted with older versions of LISREL, EQS or AMOS (Arbuckle, 1995). Special purpose programs, called MUTMUM and LINLOG, were written by Browne (1991) and
Browne and Strydom (1991), respectively. LINLOG uses non-iterative methods to fit the CDP model (see Browne and Strydom, In Press).


6. Concluding Remarks
The CDP model has many desirable properties: it provides insight into the Campbell and Fiske conditions, it has a multiplicative property suggested as appropriate for MTMM correlation matrices by Campbell and O'Connell and it can be fitted to the correlation matrix by means of conventional software.

The CDP model may not be suitable to every MTMM correlation matrix. If the model is not acceptable one may consider using an additive model instead. Presently, the confirmatory factor model is considered virtually useless for the analysis of MTMM data (Wothke, 1995, p. 9). The arguments are that: its parameters are difficult to interpret (Browne, 1984, 1995a, 1995b) and confirmatory factor analyses frequently give inadmissible or uninterpretable solutions (Brannick and Spector, 1990). Alternative additive models have been proposed by Wothke (1984, 1995), Browne (1989), Dudgeon (1994) and Kiers, Takane and Ten Berge (in press).

References


Neale, M. C. (1995). *New features in Mx 1.26*. Department of Psychiatry, Medical College of Virginia, Box 980710, Richmond, Virginia 23298, U.S.A.


Appendix: A LISREL8 setup for the CDP model with 3 traits and 3 methods.

title: CDP model on 9x9 matrix
da ni=9 ma=km ng=1 no=
km fi=
mo ly=nu,fr ne=9 ny=9 ly=di,fr ps=sy,fr te=di,fr
pa ps
0
10
110
1110
11110
111110
1111110
11111110
111111110
eq ps(2,1) ps(5,4) ps(8,7)
eq ps(3,1) ps(6,4) ps(9,7)
eq ps(3,2) ps(6,5) ps(9,8)
eq ps(4,1) ps(5,2) ps(6,3)
eq ps(7,1) ps(8,2) ps(9,3)
eq ps(7,4) ps(8,5) ps(9,6)
co ps(4,2)=ps(2,1)*ps(4,1)
co ps(5,1)=ps(2,1)*ps(4,1)
co ps(4,3)=ps(3,1)*ps(4,1)
co ps(6,1)=ps(3,1)*ps(4,1)
co ps(5,3)=ps(3,2)*ps(4,1)
co ps(6,2)=ps(3,2)*ps(4,1)
co ps(7,2)=ps(2,1)*ps(7,1)
co ps(8,1)=ps(2,1)*ps(7,1)
co ps(7,3)=ps(3,1)*ps(7,1)
co ps(9,1)=ps(3,1)*ps(7,1)
co ps(8,3)=ps(3,2)*ps(7,1)
co ps(9,2)=ps(3,2)*ps(7,1)
co ps(7,5)=ps(2,1)*ps(7,4)
co ps(8,4)=ps(2,1)*ps(7,4)
co ps(7,6)=ps(3,1)*ps(7,4)
co ps(9,4)=ps(3,1)*ps(7,4)
co ps(8,6)=ps(3,2)*ps(7,4)
co ps(9,5)=ps(3,2)*ps(7,4)
va 1 ps(1,1) ps(2,2) ps(3,3) ps(4,4) ps(5,5)
va 1 ps(6,6) ps(7,7) ps(8,8) ps(9,9)
ou

Ontvangen: 12-7-1995
Geaccepteerd: 5-3-1996