DOI:10.6129/CJP.201903_61(1).0002
中華心理學刊 民 108，61 卷，1 期，25-50
Chinese Journal of Psychology 2019, Vol.61, No.1, 25-50

The Nature of Multidimensional Constructs Represented by Item Parcels in Structural Equation Modeling

You-Lin Chen Li-Jen Weng
National Taiwan University

Item parcels, represented as summation or average over items, can be used as indicators of latent variables in structural equation modeling (SEM). Little et al. (2013) and Cole et al. (2016) indicated that the nature of the multidimensional constructs represented by parcels can be different from that assumed by the researchers. Researchers should thus examine the nature of multidimensional constructs represented by parcels prior to the analysis. Cole et al. and Williams and O'Boyle (2008) pointed out that many constructs investigated in psychological research are multidimensional, consisting of several facets, and item parcels are frequently used as indicators for these constructs in SEM. The present study therefore extended the algebraic derivation of the covariance matrix of item parcels in Sterba & MacCallum (2010) from unidimensional to multidimensional constructs to provide a theoretical framework for examining the nature of multidimensional constructs implied by parcels. The effects of parceling strategy, correlations among facets, and factorial complexity of items on the nature of multidimensional constructs represented by parcels were discussed using this framework and illustrated with a numerical simulation example. Researchers are encouraged to apply the framework proposed in this study to clarify the nature of the multidimensional constructs inferred from item parcels by examining the covariance matrix among parcels to avoid misleading results from SEM analysis.

Keywords: *item parcels, multidimensional constructs, structural equation modeling*

Summary

Item parcels, represented as a summation or average of a set of items, have become increasingly popular in structural equation modeling (SEM) applications. When the aim of an SEM analysis is to capture the relationships between latent variables rather than to investigate the relationships between items and factors, parcels can be used to enhance the reliability, continuity, and normality of the indicators for the latent constructs or to yield an optimal sample size for the analysis (e.g., Bandalos, 2002; Bandalos & Finney, 2001). Cole, Perkins, and Zelkowitz (2016) and Williams and O'Boyle (2008) pointed out that many of the constructs investigated in psychological research are multidimensional, that is, they consist of several facets, and item parcels are frequently used as indicators for these constructs in SEM. Working with multidimensional constructs, Little, Rhemtulla, Gibson, and Schoemann (2013) and Cole et al. (2016) have suggested that the parcels do not always represent the multidimensional constructs in the way the researchers assume. Thus, researchers must examine the nature of multidimensional constructs represented by parcels before conducting analysis. This study extended the algebraic derivation of the covariance matrix of item parcels in Sterba and MacCallum (2010) from unidimensional to multidimensional constructs to provide a theoretical framework for examining the nature of multidimensional constructs as represented by parcels.

Facet-representative strategy and domain-representative strategy are two major strategies for constructing item parcels with multidimensional constructs (Coffman & MacCallum, 2005; Little et al., 2013). Facet-representative parcels combine items from the same facet, whereas items from different facets are mingled to form domain-representative parcels. The facet-

representative strategy is also referred to in the literature as homogeneous parceling (Coffman & MacCallum, 2005; Cole et al., 2016), internal consistent parceling (Kishton & Widaman, 1994), or the isolated uniqueness strategy (Hall, Snell, & Foust, 1999) that items sharing the same uniqueness are pooled together to form parcels. In contrast, the domain-representative strategy combines items from different facets so that each parcel adequately represents the multiple facets in a broad domain. This parceling strategy is also known as heterogeneous parceling (Cole et al., 2016) or the distributed uniqueness strategy (Hall et al., 1999).

The higher-order factor model has been used to specify the relationships between items and multidimensional constructs (Coffman & MacCallum, 2005; Cole et al., 2016; Little et al., 2013). The first-order factors in this model represent the facets of a multidimensional construct, and the second-order factors capture the common part among the first-order factors. It should be noted that the second-order factors may or may not match the multidimensional constructs that the researchers assume. There exist two perspectives on the nature of a multidimensional construct (Cole et al., 2016; Law, Wong, & Mobley, 1998). A multidimensional construct can be conceptualized as the common part shared by different facets. Alternatively, a multidimensional construct can be hypothesized as containing not only the facets' shared aspect, but also the uniqueness of each facet. Researchers should identify the nature of the multidimensional constructs they have theorized before applying parcels in an SEM analysis and scrutinize whether the constructs represented by the parcels match their theoretical conceptualization.

The meaning of the constructs implied by the parcels depends in part on how the individual items are combined (Hall et al., 1999). Cole et al. (2016) illustrated the source of variance of a second-order factor derived from two different parceling strategies with unidimensional items. When facet-representative parcels are constructed, the second-order factors contain only the variance due to the common part shared by all the facets. In contrast, if domain-representative parcels are used, a second-order factor capture not only the variance due to the common part among all the facets, but also the specific variances of the individual first-order factors. Parceling strategy clearly affects the nature of the multidimensional constructs represented by the parcels. The algebraic derivation of parcel variance and the covariance between parcels in Little et al. (2013) also revealed the distinct natures of the multidimensional constructs extracted from the two types of parcels. In addition to the parceling strategy, the nature of a multidimensional construct represented by parcels is affected by the degree of correlations between the facets (Cole et al., 2016; Kishton & Widaman, 1994) and the factorial complexity of the items (Cole et al., 2016; Marsh, Ludtke, Nagengast, Morin, & Von Davier, 2013) as well.

Sterba and MacCallum (2010) showed algebraically that, in the population, when items are unidimensional, the covariance matrix among the factors at the item level is identical to that at the parcel level, regardless of the item-parcel allocation methods. Accordingly, in this case, the nature of the construct implied by the parcels is the same as that assumed for items. As an initial attempt to explore the nature of multidimensional constructs represented by parcels, this study was confined to the population. The algebraic derivation of the covariance matrix of item parcels in Sterba and MacCallum was extended to multidimensional constructs to provide a theoretical framework for examining the nature of the constructs implied by parcels constructed from items of multidimensional constructs. The effects of the parceling strategy, correlations between facets, and factorial complexity of the items on the nature of multidimensional constructs represented by parcels are discussed using this framework and illustrated with a numerical simulation example.

METHOD

The higher-order factor model was used to describe the relationships between the items and latent variables (Coffman & MacCallum, 2005; Cole et al., 2016; Little et al., 2013). The population covariance matrix of item parcels from multidimensional constructs was derived algebraically following Sterba and MacCallum (2010) to provide a framework for exploring the nature of multidimensional constructs represented by parcels. The effects of inter-facet correlations and factorial complexity of the items on the parcels' representations of multidimensional constructs were examined using the resulting framework.

Following the algebraic derivation of the covariance matrix of item parcels, a small-scale numerical simulation was conducted to illustrate the derivation and the effects of the parceling strategy, correlations between facets, and factorial complexity of the items on parcels' representations of constructs and the results from SEM analysis. Continuous data with a normal distribution were generated based on the model in Coffman and MacCallum (2005). Yet, in this study, all of the first-order factor loadings were fixed at 0.5, and all of the second-order factor loadings were set as identical with a range of 0.1 to 0.9 at increments of 0.1. This model represented the pure measure condition under which each item was affected by one facet. To increase the factorial complexity of part of the items, the first item of each first-order factor was specified as being affected by another facet with the dual loadings set at 0.5. The parcels formed by the two different strategies were then fitted to the structural part of the data generation model with parcels as the indicators. For each of the 18 conditions (2 factorial complexity of items × 9 second-order factor loadings), 30 replications with N of 100,000 observations were analyzed by R 3.4.0 (R core team, 2017). Note that the multidimensional constructs implied by the data generation model constituted only the common variance in facets.

RESULTS

Covariance matrix of item parcels with multidimensional constructs

Let y_i be an $m \times 1$ vector of m item scores deviated from the means, where subscript i denotes item-level quantities. Vector y_i is assumed to be indicators of q factors (F), which in turn are affected by r higher-order factors (D). The relationship between y_i and F can be represented as

$$y_i = \mathbf{\Lambda}_i \mathrm{F} + \epsilon_i, \tag{1}$$

where F is a $q \times 1$ vector of the first-order factor scores, $\mathbf{\Lambda}_i$ is an $m \times q$ matrix of the first-order factor loadings, ϵ_i is an $m \times 1$ vector of measurement errors, E(F) = 0, E(ϵ_i) = 0, and E(ϵ_iF′) =0.

In the higher-order factor model, first-order factors represent the facets embedded in a multidimensional construct. The common variance underlying the facets is delineated by the second-order factor, and the uniqueness of each facet is represented by the residual. Vector y_i can be further expressed as

$$y_i = \mathbf{\Lambda}_i \mathrm{F} + \epsilon_i = \mathbf{\Lambda}_i(\mathbf{\Gamma}_i \mathrm{D} + \zeta_i) + \epsilon_i, \tag{2}$$

where D is an $r \times 1$ vector of the second-order factor scores, $\mathbf{\Gamma}_i$ is a $q \times r$ matrix of second-order factor loadings, ζ_i is a $q \times 1$ vector of residuals, E(D) = 0, E(ζ_i) = 0, and E(Dζ_i') = 0.

Let $\mathbf{A}$ be an $n \times m$ selection matrix for selecting items in y_i to form n parcels, and $y_p = \mathbf{A}y_i$ is an $n \times 1$ vector with subscript p denoting parcel-level quantities. The population covariance matrix of the n parcels can be represented as

$$\mathrm{E}(y_p y_p') = \mathbf{\Sigma}_p = \mathbf{A\Lambda}_i\mathbf{\Gamma}_i\mathbf{\Phi\Gamma}_i'\mathbf{\Lambda}_i'\mathbf{A}' + \mathbf{A\Lambda}_i\mathbf{\Psi}_i\mathbf{\Lambda}_i'\mathbf{A}' + \mathbf{A\Theta}_{\epsilon i}\mathbf{A}', \tag{3}$$

where $\mathbf{\Phi}$ = E(DD′), $\mathbf{\Psi}_i$ = E($\zeta_i\zeta_i'$), and $\mathbf{\Theta}_{\epsilon i}$ = E($\epsilon_i\epsilon_i'$) being a diagonal matrix. As Equation (3) shows, $\mathbf{\Sigma}_p$ consists of three parts, each with a unique feature. Let **X**, **Y**, and **Z** represent the three parts of Equation (3), respectively:

$$\mathbf{X} = \mathbf{A\Lambda}_i\mathbf{\Gamma}_i\mathbf{\Phi\Gamma}_i'\mathbf{\Lambda}_i'\mathbf{A}', \tag{4}$$

$$\mathbf{Y} = \mathbf{A\Lambda}_i\mathbf{\Psi}_i\mathbf{\Lambda}_i'\mathbf{A}', \tag{5}$$

$$\mathbf{Z} = \mathbf{A\Theta}_{\epsilon i}\mathbf{A}'. \tag{6}$$

Matrix **X**, containing $\mathbf{\Phi}$, is related to the common variance shared among the facets in this representation. This common variance is always embedded in the multidimensional constructs implied by the parcels, regardless of the assumptions about the nature of the multidimensional construct. Matrix **Z**, containing $\boldsymbol{\Theta}_{\epsilon i}$, is related to the measurement errors of the items and does not affect the nature of the construct represented by the parcels in this framework.

Matrix **Y**, containing $\boldsymbol{\Psi}_i$, is related to the uniqueness variances of the facets in this representation. The nature of a multidimensional construct implied by the parcels depends on matrix **Y**. In particular, the characteristics of the diagonal and the offdiagonal elements under every second-order factor simultaneously influence the nature of the represented multidimensional construct. If the elements in the diagonal share no common components and the off-diagonal elements are all 0, the uniquenesses in the facets are considered as measurement errors of the parcels, resulting in a multidimensional construct that simply represents the common variance among the facets. In contrast, if some of the off-diagonal elements are non-zero, the diagonal elements would share certain variation from the uniquenesses of the facets, resulting in a multidimensional construct that includes not only the common variance among the facets but also the uniquenesses of the facets.

As can be seen from Equation (5), the feature of **Y** is simultaneously affected by the parceling strategy (**A**), the correlations between the facets as reflected by the second-order factor loadings ($\boldsymbol{\Gamma}_i$), and the factorial complexity of the items ($\boldsymbol{\Lambda}_i$). The higher the inter-facet correlations, the larger the second-order factor loadings, and the smaller the uniqueness variances of the facets ($\boldsymbol{\Psi}_i$). The magnitude of the uniqueness variances embedded in a multidimensional construct represented by parcels is thus also contingent on the correlations between facets. When the facets are highly correlated, the multidimensional constructs represented by the parcels would mostly reflect the common variance shared by the facets.

Example with artificial data

The current data generation model assumes that a multidimensional construct contains only the source of variation common to all of its facets. Thus, based on the above derivation, the nature of the multidimensional constructs represented by the parcels will deviate from the constructs assumed by the data generation model except when the facet-representative parcels are constructed from pure measures. As expected, the parameter estimates obtained from the facet-representative parcels with pure measures are less biased than the others. Moreover, when part of the items are complex, that is, affected by two facets, the parameter estimates are biased regardless of the parceling strategy used, as the implied multidimensional constructs contain not only the common variance but also the uniquenesses of the facets. The bias of estimation decreases as the inter-facet correlations increase, as signified by the rising second-order factor loadings. The higher the second-order factor loadings, the smaller the magnitude of the uniquenesses contained in the multidimensional constructs. Thus, the results of the SEM analysis are affected by the parceling strategy, inter-facet correlations, and item factorial complexity.

DISCUSSION

Parceling should only be used in SEM when the analysis focuses on the relationships between the latent variables. If the relationships between items and constructs are of interest, as in scale development, the items should not be parceled. The nature of the multidimensional constructs represented by item parcels are affected not only by the parceling strategy, but also by the correlations between the facets of the broad domain, and the factorial complexity of the items. As an initial attempt to explore the nature of the constructs implied by item parcels from multidimensional constructs, the current algebraic derivation is confined to the population level. Future studies should take sampling variability into account. Comprehensive simulations are also called for to fully understand the effects of the parceling strategy, inter-facet correlations, and factorial complexity of the items on the results of SEM analysis. This study provides a preliminary basis for understanding the meaning of constructs represented by item parcels through the decomposition of the covariance matrix of parcels. Researchers are encouraged to apply the proposed framework to item parcels to uncover the nature of the constructs they infer.