Why the Items versus Parcels Controversy Needn’t Be One

Todd D. Little*
University of Kansas
Director, Quantitative Psychology Training Program
Director, Center for Research Methods and Data Analysis
Director, Undergraduate Social and Behavioral Sciences Methodology Minor
Member, Developmental Psychology Training Program

crmda.KU.edu

Colloquium presented 10-04-2012 @ Notre Dame University
Based on my Presidential Address presented 08-04-2011 @
American Psychological Association Meeting in Washington, DC

*see slide 2
Key Sources and Acknowledgements

Special thanks to:

Mijke Rhemtulla, Kimberly Gibson, Alex Schoemann, Wil Cunningham, Golan Shahar, Keith Widaman, John Graham, Ulman Lindenberger, & John Nesselroade

Based on:

• Little, T. D., Rhemtulla, M., Gibson, K., & Schoemann, A. M. (in press). Why the items versus parcels controversy needn’t be one. Psychological Methods, 00, 000-000.


• Little, T. D., Lindenberger, U., & Nesselroade, J. R. (1999). On selecting indicators for multivariate measurement and modeling with latent variables: When "good" indicators are bad and "bad" indicators are good. Psychological Methods, 4, 192-211.

Overview

• Learn what parcels are and how to make them
• Learn the reasons for and the conditions under which parcels are beneficial
• Learn the conditions under which parcels can be problematic
• Disclaimer: This talk reflects my view that parcels per se aren’t controversial if done thoughtfully.
What is Parceling?

• Parceling: Averaging (or summing) two or more items to create more reliable indicators of a construct

• ≈ Packaging items, tying them together

• Data pre-processing strategy
A CFA of Items

Model Fit: $\chi^2_{(53, n=759)} = 181.2$; RMSEA = 0.056 (.048-.066); NNFI/TLI = .97; CFI = .98
CFA: Using Parcels

Average 2 items to create 3 parcels per construct

(6.2.Parcels)
CFA: Using Parcels

Similar solution
- Similar factor correlation
- Higher loadings, more reliable info
- Good model fit, improved $\chi^2$

Model Fit: $\chi^2_{(8, n=759)} = 26.8$; RMSEA = $0.056_{(0.033-0.079)}$; NNFI/TLI = $0.99$; CFI = $0.99$
Philosophical Issues

To parcel, or not to parcel…?
Pragmatic View

“Given that measurement is a strict, rule-bound system that is defined, followed, and reported by the investigator, the level of aggregation used to represent the measurement process is a matter of choice and justification on the part of the investigator”

Preferred terms: remove unwanted, clean, reduce, minimize, strengthen, etc.

From Little et al., 2002
Empiricist / Conservative View

“Parceling is akin to cheating because modeled data should be as close to the response of the individual as possible in order to avoid the potential imposition, or arbitrary manufacturing of a false structure”

Preferred terms: mask, conceal, camouflage, hide, disguise, cover-up, etc.

From Little et al., 2002
Psuedo-Hobbesian View

Parcels should be avoided because researchers are ignorant (perhaps stupid) and prone to mistakes. And, because the unthoughtful or unaware application of parcels by unwitting researchers can lead to bias, they should be avoided.

Preferred terms: most (all) researchers are un___ as in … unaware, unable, unwitting, uninformed, unscrupulous, etc.
Other Issues I

- Classical school vs. Modeling School
  - Objectivity versus Transparency
  - Items vs. Indicators
  - Factors vs. Constructs

- Self-correcting nature of science

- Suboptimal simulations
  - Don’t include population misfit
  - Emphasize the ‘straw conditions’ and proofing the obvious; sometimes over generalize
Other Issues II

• Focus of inquiry
  – Question about the items/scale development?
    • Avoid parcels
  – Question about the constructs?
    • Parcels are warranted but must be done thoughtfully!
  – Question about factorial invariance?
    • Parcels are OK if done thoughtfully.
Measurement

“Whatever exists at all exists in some amount. To know it thoroughly involves knowing its quantity as well as its quality”
- E. L. Thorndike (1918)

- Measurement starts with **Operationalization**
  Defining a concept with specific observable characteristics
  [Hitting and kicking ~ operational definition of Overt Aggression]

- Process of linking constructs to their manifest *indicants*
  (object/event that can be seen, touched, or otherwise recorded; cf. items vs. indicators)

- Rule-bound assignment of numbers to the *indicants* of that which exists
  [e.g., Never=1, Seldom=2, Often=3, Always=4]

- … although convention often ‘rules’, the rules should be chosen and defined by the investigator
“Indicators are our worldly window into the latent space”

- John R. Nesselroade
Classical Variants

a) \( X_i = T_i + S_i + e_i \)
b) \( X = T + S + e \)
c) \( X_1 = T_1 + S_1 + e_1 \)

\( X_i \) : a person’s observed score on an item
\( T_i \) : 'true' score (i.e., what we hope to measure)
\( S_i \) : item-specific, yet reliable, component
\( e_i \) : random error or noise.

Assume:
• \( S_i \) and \( e_i \) are normally distributed (with mean of zero) and uncorrelated with each other
• Across all items in a domain, the \( S_i \)s are uncorrelated with each other, as are the \( e_i \)s
Latent Variable Variants

\[ a) \ X_1 = T + S_1 + e_1 \]
\[ b) \ X_2 = T + S_2 + e_2 \]
\[ c) \ X_3 = T + S_3 + e_3 \]

\( X_1 - X_3 \) : are multiple indicators of the same construct
\( T \) : common 'true' score across indicators
\( S_1 - S_3 \) : item-specific, yet reliable, component
\( e_1 - e_3 \) : random error or noise.

Assume:
• \( S_s \) and \( e_s \) are normally distributed (with mean of zero) and uncorrelated with each other
• Across all items in a domain, the \( S_s \) are uncorrelated with each other, as are the \( e_s \)
Construct = Common Variance of Indicators
Construct = Common Variance of Indicators
Empirical Pros

Psychometric Characteristics of Parcels (vs. Items)

- Higher reliability, communality, & ratio of common-to-unique factor variance
- Lower likelihood of distributional violations
- More, Smaller, and more-equal intervals

<table>
<thead>
<tr>
<th></th>
<th>Never</th>
<th>Seldom</th>
<th>Often</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happy</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Glad</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Mean</td>
<td>1</td>
<td>1.5</td>
<td>2</td>
<td>2.5</td>
</tr>
<tr>
<td>Sum</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
More Empirical Pros

Model Estimation and Fit with Parcels (vs. Items)

• Fewer parameter estimates
• Lower indicator-to-subject ratio
• Reduces sources of parsimony error (population misfit of a model)
  ▪ Lower likelihood of correlated residuals & dual factor loading
• Reduces sources of sampling error
• Makes large models tractable/estimable
Sources of Variance in Items
Simple Parcel

\[
\text{var}\left( \frac{(I_2 + I_3 + I_4)}{3} \right) = \text{var}(T_1) + \frac{1}{9} \left[ \text{var}(s_2) + \text{var}(s_3) + \text{var}(s_4) + \text{var}(x) \right] + \text{var}(e_2) + \text{var}(e_3) + \text{var}(e_4)
\]

\[
T = s_2 + s_3 + s_4 + e_2 + e_3 + e_4 = 3
\]

\[
T = 1/9 \text{ of their original size!}
\]

\[
\begin{align*}
U_3 & \quad \text{U}_1 & \quad C & \quad \text{U}_2 \\
\text{T}_3 & \quad \text{T}_1 & \quad \text{T}_2 & \\
\text{I}_1 & \quad \text{I}_2 & \quad \text{I}_3 & \quad \text{I}_4 & \quad \text{I}_5 & \quad \text{I}_6
\end{align*}
\]

\[
\begin{align*}
\text{s}_1 + \text{e}_1 & \quad \text{s}_2 + \text{e}_2 & \quad \text{s}_3 + \text{e}_3 & \quad \text{s}_4 + \text{e}_4 & \quad \text{s}_5 + \text{e}_5 & \quad \text{s}_6 + \text{e}_6 \\
\text{s}_4 & \quad \text{x}
\end{align*}
\]
Correlated Residual

$$\text{var} \left( \frac{(I_3 + I_4 + I_5)}{3} \right) = \text{var}(T_1) + \frac{4}{9} \text{var}(x) + \frac{1}{9} \left[ \text{var}(s_3) + \text{var}(s_4) + \text{var}(s_5) + \text{var}(e_3) + \text{var}(e_4) + \text{var}(e_5) \right]$$

$$3 = \frac{4}{9} \text{its original size!}$$

$$\frac{1}{9} \text{of their original size!}$$
Cross-loading: Correlated Factors

\[
\text{var}\left(\frac{(I_2 + I_3 + I_6)}{3}\right) = \text{var}(U_1) + \frac{16}{9} \text{var}(C) + \frac{1}{9} \left[ \text{var}(U_2) + \text{var}(s_2) + \text{var}(s_3) + \text{var}(s_6) + \text{var}(e_2) + \text{var}(e_3) + \text{var}(e_6) \right]
\]

\[
\begin{bmatrix}
U_1 \\
C \\
s_2 \\
e_2 \\
\end{bmatrix} + 
\begin{bmatrix}
U_1 \\
C \\
s_3 \\
e_3 \\
\end{bmatrix} + 
\begin{bmatrix}
U_1 \\
C \\
s_5 \\
e_5 \\
\end{bmatrix} = 3
\]

16/9 its original size!

1/9 of their original size!
Cross-loading: Uncorrelated Factors

\[
\begin{align*}
\text{var}\left(\frac{(I_1 + I_2 + I_3)}{3}\right) &= \text{var}(U_1) + \text{var}(C) + \frac{1}{9}\left[\text{var}(U_3) + \text{var}(s_1) + \text{var}(s_2) + \text{var}(s_3) + \text{var}(e_1) + \text{var}(e_2) + \text{var}(e_3)\right]
\end{align*}
\]

\[
\begin{align*}
\begin{array}{c}
\text{U}_1 \\
\text{C} \\
\text{s}_2 \\
\text{e}_2
\end{array} + 
\begin{array}{c}
\text{U}_1 \\
\text{C} \\
\text{s}_3 \\
\text{e}_3
\end{array} + 
\begin{array}{c}
\text{U}_1 \\
\text{C} \\
\text{s}_1 \\
\text{e}_1
\end{array}
\end{align*}
\]

\[
\begin{array}{c}
\text{T}_1 \\
\text{T}_3
\end{array}
\]

\[
\begin{array}{c}
\text{U}_1 \\
\text{C}
\end{array}
\]

= \frac{1}{9} \text{ of their original size!}
Empirical Cautions*

*Shared method variance isn’t fixed with parcels
Construct = Common Variance of Indicators
Construct = Common Variance of Indicators
Three is Ideal

- Positive 1
  - Great & Glad
    - θ_{11}
  - Cheerful & Good
    - θ_{22}
  - Happy & Super
    - θ_{33}
  - Terrible & Sad
    - θ_{44}
  - Down & Blue
    - θ_{55}
  - Unhappy & Bad
    - θ_{66}

- Negative 2
  - Ψ_{21}
  - Ψ_{22}
  - Ψ_{11}
  - Ψ_{12}
Three is Ideal

Matrix Algebra Formula: $\Sigma = \Lambda \Psi \Lambda' + \Theta$

<table>
<thead>
<tr>
<th></th>
<th>Great &amp; Glad</th>
<th>Cheerful &amp; Good</th>
<th>Happy &amp; Super</th>
<th>Terrible &amp; Sad</th>
<th>Down &amp; Blue</th>
<th>Unhappy &amp; Bad</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>$\lambda_{11} \psi_{11} \lambda_{11} + \theta_{11}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td>$\lambda_{11} \psi_{11} \lambda_{21}$</td>
<td>$\lambda_{21} \psi_{11} \lambda_{21} + \theta_{22}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P3</td>
<td>$\lambda_{11} \psi_{11} \lambda_{31}$</td>
<td>$\lambda_{21} \psi_{11} \lambda_{31}$</td>
<td>$\lambda_{31} \psi_{11} \lambda_{31} + \theta_{33}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N1</td>
<td>$\lambda_{11} \psi_{11} \lambda_{42}$</td>
<td>$\lambda_{21} \psi_{21} \lambda_{42}$</td>
<td>$\lambda_{31} \psi_{21} \lambda_{42}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2</td>
<td>$\lambda_{11} \psi_{11} \lambda_{52}$</td>
<td>$\lambda_{21} \psi_{21} \lambda_{52}$</td>
<td>$\lambda_{31} \psi_{21} \lambda_{52}$</td>
<td>$\lambda_{42} \psi_{22} \lambda_{52} + \theta_{44}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N3</td>
<td>$\lambda_{11} \psi_{11} \lambda_{62}$</td>
<td>$\lambda_{21} \psi_{21} \lambda_{62}$</td>
<td>$\lambda_{31} \psi_{21} \lambda_{62}$</td>
<td>$\lambda_{42} \psi_{22} \lambda_{62}$</td>
<td>$\lambda_{52} \psi_{22} \lambda_{52} + \theta_{55}$</td>
<td>$\lambda_{62} \psi_{22} \lambda_{62} + \theta_{66}$</td>
</tr>
</tbody>
</table>

- Cross-construct item associations (in box) estimated only via $\Psi_{21}$ – the latent constructs’ correlation.
- Degrees of freedom only arise from between construct relations.
Empirical Cons

• **Multidimensionality**
  • Constructs and relationships can be hard to interpret if done improperly

• **Model misspecification**
  • Can get improved model fit, regardless of whether model is correctly specified
  • Increased Type II error rate if question is about the items

• **Parcel-allocation variability**
  • Solutions depend on the parcel allocation combination (Sterba & McCallum, 2010; Sterba, in press)
    • Applicable when the conditions for sampling error are high
Psychometric Issues

• Principles of Aggregation (e.g., Rushton et al.)
  • Any one item is less representative than the average of many items (selection rationale)
  • Aggregating items yields greater precision

• Law of Large Numbers
  • More is better, yielding more precise estimates of parameters (and a person’s true score)
  • Normalizing tendency
Construct Space with Centroid
Potential Indicators of the Construct
Selecting Six (Three Pairs)
... take the mean
... and find the centroid
Selecting Six (Three Pairs)
... take the mean
… find the centroid
How about 3 sets of 3?
… taking the means
... yields more reliable & accurate indicators
Building Parcels

• Theory – Know thy S and the nature of your items.
• Random assignment of items to parcels (e.g., fMRI)
  • Use Sterba’s calculator to find allocation variability when sampling error is high.
• Balancing technique
  • Combine items with higher loadings with items having smaller loadings [Reverse serpentine pattern]
• Using a priori designs (e.g., CAMI)
  • Develop new tests or measures with parcels as the goal for use in research
Techniques: Multidimensional Case

Example: ‘Intelligence’ ~ Spatial, Verbal, Numerical

• **Domain Representative Parcels**
  • Has mixed item content from various dimensions
  • Parcel consists of: 1 Spatial item, 1 Verbal item, and 1 Numerical item

• **Facet Representative Parcels**
  • Internally consistent, each parcel is a ‘facet’ or singular dimension of the construct
  • Parcel consists of: 3 Spatial items
  • Recommended method
Domain Representative Parcels

Spatial + Verbal + Numerical = Parcel #1

Spatial + Verbal + Numerical = Parcel #2

Spatial + Verbal + Numerical = Parcel #3
Domain Representative
But which facet is driving the correlation among constructs?
Facet Representative Parcels

Parcel: Spatial

Parcel: Verbal

Parcel: Numerical
Facet Representative
Facet Representative

Intellective Ability

Diagram depicts smaller communalities (amount of shared variance)
Facet Representative Parcels

A more realistic case with higher communalities
Facet Representative
Facet Representative

Intellective Ability

Parcels have more reliable information
2nd Order Representation

Capture multiple sources of variance?
2nd Order Representation

Variance can be partitioned even further
2nd Order Representation

Lower-order constructs retain facet-specific variance
Functionally Equivalent Models

Explicit Higher-Order Structure

Implicit Higher-Order Structure
When Facet Representative Is Best

A) A higher-order representation of three related constructs

B) An equivalent facet-representative parcel-based version
When Domain Representative Is Best

A) A higher-order representation of three related constructs

B) An equivalent domain-representative parcel-based version
Thank You!

Todd D. Little
University of Kansas

Director, Quantitative Psychology Training Program
Director, Center for Research Methods and Data Analysis
Director, Undergraduate Social and Behavioral Sciences Methodology Minor
Member, Developmental Psychology Training Program

crmda.KU.edu