Picturing the impossible: Pictures, repetitions, and item-set card effects in measuring vocabulary

Eleanor Fang Yan & Lee Branum-Martin
Georgia State University
Agenda

• Introduction
• Assessment Design
• Statistical Approaches
• Results
• Conclusions
• Limitations & Future Directions
Introduction

• Background
  • With reliable design and valid scoring, subtests of language assessment batteries are expected to show good global fit of a single factor confirmatory model.
  • TOLD-4: Test of Language Development, 4th edition
    • IPV: Intermediate level of Picture Vocabulary Subtest
    • IPV’s complicated assessment design features make a single-scale model impossible.

• Purpose
  • To examine the psychometric properties and structural validity of a picture vocabulary test
Assessment Design

• TOLD-4, IPV
  • 9 picture cards
  • Each with 6 pictures
  • 7 – 11 items per card
  • 80 items in total
  • Ceiling rule
    • Stop testing when participants miss 3 consecutive items.
Picture Card 1

1. tail wagger
2. hibernating mammal
3. branch swinger
4. bird chaser
5. grizzled carnivore
6. abrasive surface
7. playful primate
8. keen edged
9. ursus posing
## Assessment Design (cont.)

<table>
<thead>
<tr>
<th>Card</th>
<th>Item Details</th>
<th>Items that share the same answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9 : 1-9</td>
<td>1,4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2,5,9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3,7</td>
</tr>
<tr>
<td>2</td>
<td>11 : 10-20</td>
<td>10,20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12,14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15,19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16,17,18</td>
</tr>
<tr>
<td>3</td>
<td>8 : 21-28</td>
<td>21,26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22,27,28</td>
</tr>
<tr>
<td>4</td>
<td>10 : 29-38</td>
<td>30,38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>31,33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>32,35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>34,37</td>
</tr>
<tr>
<td>5</td>
<td>8 : 39-46</td>
<td>39,42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40,45</td>
</tr>
<tr>
<td>6</td>
<td>7 : 47-53</td>
<td>48,50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>49,53</td>
</tr>
<tr>
<td>7</td>
<td>8 : 54-61</td>
<td>55,61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>57,58</td>
</tr>
<tr>
<td>8</td>
<td>10 : 62-71</td>
<td>62,69,70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>64,67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>68,71</td>
</tr>
<tr>
<td>9</td>
<td>9 : 72-80</td>
<td>72,78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>73,80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>74,79</td>
</tr>
</tbody>
</table>
Assessment Design (cont.)

• Unique Features

  • 6 choices per item → nominal vs. binary
  • method effects of card splits → single-factor vs. bi-factor
  • repeated item answers → residual covariances
Statistical Approach

• *Ideal Model*: representing both trait and method features
  
  • Bi-factor Multifactor Nominal CFA with residual covariances
• *Ideal Model*: representing both trait and method features
  
  • Bi-factor Multifactor *Nominal* CFA with residual covariances

**Impossible!!!**
Statistical Approach (cont.)

• Single-factor CFA
  • w/ | w/o Residual Covariances

• Multifactor CFA
  • w/ | w/o Residual Covariances

• Bi-factor Multifactor CFA
  • w/ | w/o Residual Covariances

Trait-only Model

Method-only Model

Trait-Method Model
<table>
<thead>
<tr>
<th>Trait/Method</th>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>CFI</th>
<th>TLI</th>
<th>RMSEA</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trait</td>
<td>Single</td>
<td>6185</td>
<td>3002</td>
<td>.68</td>
<td>.67</td>
<td>.05</td>
<td>High Correlations &amp; Zero Variances</td>
</tr>
<tr>
<td>Trait</td>
<td>Single, residual</td>
<td>6139</td>
<td>2977</td>
<td>.68</td>
<td>.67</td>
<td>.05</td>
<td>Negative Thetas</td>
</tr>
<tr>
<td>Method</td>
<td>Nine cards</td>
<td>3433</td>
<td>2813</td>
<td>.92</td>
<td>.92</td>
<td>.02</td>
<td>Negative Thetas</td>
</tr>
<tr>
<td>Method</td>
<td>Nine cards, residual</td>
<td>2897</td>
<td>2499</td>
<td>.93</td>
<td>.93</td>
<td>.02</td>
<td>Negative Thetas</td>
</tr>
<tr>
<td>Trait-Method</td>
<td>Bifactor</td>
<td>2781</td>
<td>2476</td>
<td>.96</td>
<td>.96</td>
<td>.02</td>
<td></td>
</tr>
<tr>
<td>Trait-Method</td>
<td>Bifactor, residual</td>
<td>2519</td>
<td>2358</td>
<td>.98</td>
<td>.98</td>
<td>.01</td>
<td>Negative Thetas</td>
</tr>
</tbody>
</table>
Results (cont.) - Scatterplot of Parameters
Conclusion

• Unique Features
  • 6 choices per item
  • Method effects of card splits
  • Repeated item answers
  • Nominal vs. binary
  • Single-factor vs. bi-factor
  • Residual covariances

TOLD-4 IPV shows a problematic structure.
Limitation/Future Directions

• Non-parametric Item Response Theory (IRT)
• Dynamic Fit Indices (DFI)
• *Looking forward to suggestions!*
Acknowledgement

• National Institutes of Health, National Institute on Deafness and Other Communication Disorders (NIDCD)
  • Secondary Data Analysis: R21DC019732 (PI: Branum-Martin & Washington)

• National Institute of Child Health and Human Development (NICHD)
  • Original Data Collection: R24HD075454 (PI: Julie A. Washington)

• We appreciate the great work of research assistants in the lab!
  • Stephanie Diaz & Roula Aldib
Picturing the impossible:
Pictures, repetitions, and item-set card effects in measuring vocabulary

Thank you!
Eleanor Fang Yan: fyan2@student.gsu.edu
Results (cont.) - Scatterplots of Parameters
Results (cont.) - Scatterplots of Parameters
Single-Factor Model
Multi-method Model
Multi-trait Multi-method Model

General Factor

Card1

Card2

Card3

Card4

Card5

Card6

Card7

Card8

Card9

θ_{1-9}

θ_{11}

λ_{102-180}

ψ_{11}

1, λ_{202-209}

1, λ_{210-228}

1, λ_{229-238}

1, λ_{239-246}

1, λ_{247-253}

1, λ_{254-261}

1, λ_{262-271}

1, λ_{272-280}

Y1-9

Y10-20

Y21-28

Y29-38

Y39-46

Y47-53

Y54-61

Y62-71

Y72-80

θ_{17}

θ_{25/28/59}

θ_{1-20}

θ_{11-28}

θ_{29-38}

θ_{39-46}

θ_{47-53}

θ_{54-61}

θ_{62-71}

θ_{72-80}

1, λ_{202-209}

1, λ_{210-228}

1, λ_{229-238}

1, λ_{239-246}

1, λ_{247-253}

1, λ_{254-261}

1, λ_{262-271}

1, λ_{272-280}