IDENTIFYING SOURCES OF DIFFERENTIAL ITEM FUNCTIONING ON AN ENGLISH LANGUAGE PROFICIENCY ASSESSMENT

Amy Clark, Neal Kingston

*Educational Psychology and Research*
*The University of Kansas*

### METHODS

**Participants**

<table>
<thead>
<tr>
<th>Grade</th>
<th>2nd Grade</th>
<th>3rd Grade</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish</td>
<td>16,017</td>
<td>15,359</td>
<td>31,376</td>
</tr>
<tr>
<td>Vietnamese</td>
<td>694</td>
<td>611</td>
<td>1,305</td>
</tr>
<tr>
<td>Total</td>
<td>16,711</td>
<td>15,970</td>
<td>32,681</td>
</tr>
</tbody>
</table>

**Test Material**

- Kansas English Language Proficiency Assessment (Peyton et al., 2007)
  - Reading subtest had most items and was self-administered
  - Administered 2007–2011
  - Two forms (Form A – 22 items, Form B 23 items)

**Procedure**

1. Items coded for item characteristics
   - Cognates, words with suffixes, multisyllabic words, unique to English sounds from Vietnamese and Spanish (Tang, 2006; Chen & Henning, 1985)
2. Logistic regression used to analyze 45 items for uniform and nonuniform DIF
3. Effect size variable created using change in pseudo $R^2$
4. Correlation between item characteristics & effect size
5. Linear regression with significant characteristics

<table>
<thead>
<tr>
<th>Effect size</th>
<th>Multi-collinearity</th>
<th>Spanish sounds</th>
<th>Vietnamese sounds</th>
<th>Cognates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Form A Item 5</td>
<td>0.30***</td>
<td>0.23**</td>
<td>0.25**</td>
<td>0.27**</td>
</tr>
<tr>
<td>Form A Item 8</td>
<td>0.30***</td>
<td>0.23**</td>
<td>0.25**</td>
<td>0.27**</td>
</tr>
<tr>
<td>Form B Item 11</td>
<td>0.30***</td>
<td>0.23**</td>
<td>0.25**</td>
<td>0.27**</td>
</tr>
<tr>
<td>Form B Item 12</td>
<td>0.30***</td>
<td>0.23**</td>
<td>0.25**</td>
<td>0.27**</td>
</tr>
</tbody>
</table>

### RESULTS

**Uniform DIF**

- 23 total items with significant group term after controlling for proficiency
- 12 items favored Vietnamese-speaking students

**Nonuniform DIF**

- 11 total items with significant interaction term after controlling for proficiency and group
- 10 items favored Vietnamese-speaking students

Correlations between Item Characteristics and Effect Size

- Post Hoc Logistic Regression
  - Determine how well item characteristics predicted the occurrence of uniform DIF on an item
  - Classification Accuracy:
    - No predictors: 51%
    - 3 item characteristics: 78%
    - Sounds only: 82%

### CONCLUSIONS

**Important Findings**

- Presence of nonuniform DIF
- Discrepancies between equated forms (11 vs. 16 items flagged)

**Unexpected Outcomes**

- Negative correlations between effect sizes and item characteristics
- Suffixes and cognates not significant predictors of DIF by group

**Limitations**

- Miniscule effect size changes (maximum .004)
- Skewed item characteristics

**Future Research**

- Additional language groups, subtests, grade bands
- External proficiency measure
- Use of item response theory or Mantel-Haenszel method

---


