Criterion-Related Validity of Subscores in High School Diploma Examinations

Okan Bulut
Centre for Research in Applied Measurement and Evaluation
University of Alberta

Vancouver, BC – July 2, 2016
Benefits of Reporting Subscores

- A subscore refers to a score on one test in a test battery or one section of a single test.
  - For example, subscores in algebra, geometry, and arithmetic in a mathematics assessment

- Subscores can provide more diagnostically useful information than total test scores on examinees’ strengths and weaknesses for
  - Determining content areas that need instructional improvement,
  - Monitoring a student’s progress in different content areas, and
  - Identifying potential learning difficulties.
Total Score vs. Subscores

Continuous Testing: Psychometric Challenges and Opportunities
Characteristics of Subscores

- **Reliability of subscores**
  - Subscore augmentation (Wainer et al., 2001)
  - Within-person/between-person reliability (Bulut, 2013)

- **Added value of subscores over total score**
  - PRMSE (Haberman, 2008)
  - Subscore utility index (Brennan, 2012)

- **Criterion-related validity of subscores**
  - Criterion-related profile analysis (Davison et al., 2015)
Characteristics of Subscores

- Reliability of subscores
  - Subscore augmentation (Wainer et al., 2001)
  - Within-person/between-person reliability (Bulut, 2013)

- Added value of subscores over total score
  - PRMSE (Haberman, 2008)
  - Subscore utility index (Brennan, 2012)

- Criterion-related validity of subscores
  - Criterion-related profile analysis (Davison et al., 2015)
Purpose of This Study

- This study aims to investigate the criterion-related validity of subscores in a high school diploma examination.

- Criterion variables:
  - Overall achievement in the high school diploma examination
  - GPA

- Existing criterion variables ➔ Concurrent validity
Data Source

- 2013/2014 High School Diploma Exams in the province of Alberta

- January & June test administrations (N=53,418)

- Major content domains:
  - Mathematics
    - Relations and Functions
    - Trigonometry, and Permutations
    - Combinations and Binomial Theorem
  - English Language Arts
  - Social Studies
Methodology

- Criterion-related profile analysis (Davison et al., 2015)

\[ Y_p = b \]

- \( Y_p \) is the criterion variable for person \( p \)
- \( b_0 \) is the intercept
- \( b_1 \) to \( b_V \) are the regression coefficients for the subscores \( (\dots) \) in the prediction of the criterion variable
- \( \varepsilon_p \) is the residual for person \( p \)
Methodology

\[ Y_p = b_0 + \bar{b} \text{TotalScore}_p + V \text{Cov}_p + \epsilon_p \]

- \( \bar{b} \) is the mean of the regression coefficient from Step 1
- \( \text{Cov}_p = \left( \frac{1}{V} \right) \sum_{v=1}^{V} (b_v - \bar{b})(\text{Subscore}_p^{sv} - \text{AverageSubscore}_p) \)

The two regression equations have the same \( df \) and \( R^2 \).
Methodology

- Model comparison:
  - Reduced model (R): Total score as a sole predictor
  - Full model (F): Total score & patterns of subscores as predictors

If significant, subscores don’t explain any variance beyond the total score

If NOT significant, patterns among subscores explain additional variance beyond the total score
Table 1.

Results of Criterion-Related Profile Analysis for Overall Achievement

<table>
<thead>
<tr>
<th>Test Administration</th>
<th>$R^2$ Full Model</th>
<th>$R^2$ Reduced Model</th>
<th>$\Delta R^2$</th>
<th>$F$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>.951</td>
<td>.936</td>
<td>.015</td>
<td>.578</td>
<td>.321</td>
</tr>
<tr>
<td>June</td>
<td>.868</td>
<td>.882</td>
<td>.014</td>
<td>.526</td>
<td>.283</td>
</tr>
</tbody>
</table>
Table 2.

Results of Criterion-Related Profile Analysis for GPA

<table>
<thead>
<tr>
<th>Test Administration</th>
<th>$R^2$ Full Model</th>
<th>$R^2$ Reduced Model</th>
<th>$\Delta R^2$</th>
<th>$F$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>.828</td>
<td>.725</td>
<td>.103</td>
<td>3.869</td>
<td>.004</td>
</tr>
<tr>
<td>June</td>
<td>.780</td>
<td>.693</td>
<td>.087</td>
<td>3.268</td>
<td>.011</td>
</tr>
</tbody>
</table>

Subscores seem to explain additional variance beyond the total score.
Conclusions

1. Total scores seem to be a stronger predictor of the overall achievement in the diploma examinations (as expected).

2. However, the subscores can explain more variation among high school students’ GPAs than the total score.
Conclusions

3. Follow-up tests with Math total scores instead of three Math subscores yielded very similar results.
   - The utility of math subscores is questionable.

4. Further studies can be conducted to examine predictive validity of the subscores in addition to concurrent validity.
Thank you!

For further information please contact:
Okan Bulut
bulut@ualberta.ca