

Item level psychometric characteristics for English Language Learners

Salih Binici

Florida Department of Education

Introduction

- In evaluating the validity of a test for a subgroup, the primary concern is test fairness, and the measure of fairness is comparability (Willingham et al., 1988).
- In situations where individuals who take accommodated versions of tests are compared with those who take the standard version, an additional validity issue is the comparability of scores across the different test formats (Sireci, 2006).

Introduction

- There are several aspects of comparability, such as test content, psychometrics, accommodations, predictive validity, etc.
- Psychometrics (Test Level vs. Item Level)
 - a) Reliability
 - b) Factor structure
 - c) Differential Item Functioning

Introduction

- Differential item functioning (DIF) is present when the probability of correct response for an item differs between the examinees, given the same level of ability.
- DIF analyses are conducted on middle school grades reading test items.
- Approximately 45 MC items in each grade level.

Method

- Uniform DIF is investigated via logistic regression
- Scale scores were used as matching criterion
- ELL vs. standard curriculum students
- ELL with accommodations vs. ELL without accommodations

ELL Accommodations

- Type of Accommodations
 1. Extended Time
 2. Flexible Scheduling
 3. Setting
 4. Assistance in Heritage Language
 5. Approved Dictionary
- Approximately 68% of proficient ELL students received some type of accommodations.

ELL Accommodations

- The primary accommodations are Setting (taking test at a different classroom) and Dictionary.
 - Unless ELL students take the test at a different classroom, it is hard to give other accommodations such as Extended Time and Flexible Scheduling except Dictionary.
 - If a dictionary is given, students are trained how to use it prior to test.
- Assistance in ELL students' native language is the least common accommodation

ELL vs. standard curriculum students

- Approximately 6% of the items displayed large DIF in each grade level.
- 4% of the items out of 6% exhibited DIF favoring ELL students.
- The remaining 2% of the items exhibited DIF favoring the standard curriculum students.

ELL with accommodations vs. ELL without accommodations

- About 8% of the tested items had large DIF magnitudes.
- 5% of these items were advantageous for the ones with accommodations.
- The remaining 3% favored the ones without accommodations.

Issues to consider next

- Review of these items by the content experts.
- Conducting DIF analyses considering the specific type of accommodation and the number of accommodations.
- Dividing the ELL students into categories, such as Proficient vs. Non-Proficient based on the English language proficiency indicators.