Multi-Stage Computerized Adaptive Testing: Using Item Response Theory for Design Selection

Michael C. Edwards and David Thissen
L.L. Thurstone Psychometric Laboratory
University of North Carolina at Chapel Hill

CAT Designs
- "Typical" CAT
  - Item by item adaptation
  - Content specifications (on the fly)
  - Exposure control (on the fly)
- Multi-form Structure (MFS) CAT
  - Block by block adaptation
  - Pre-assembled blocks
  - Content specifications met in assembly
  - Human review of forms

Uniform Multi-form Structure (uMFS) CAT
- uMFS-CAT allows us to:
  - Have uniform item exposure by construction, controlling "absolute exposure for expected candidate volumes"
- CAST, BMAT (Luecht and colleagues)

What does a uMFS-CAT look like?

Goals for uMFS-CAT (Objective Function)
1. Small average error variance (1-\(\rho\))
2. Uniform accuracy of theta measurement
3. Routing stage blocks equivalent
4. Branching fractions as specified

Questions we addressed for the MCAT
- Given a set of reasonable possibilities:
  - What "shape" of uMFS might work best?
  - How should items be arranged within that shape?
  - What are the psychometric properties of a uMFS computerized adaptive MCAT?
Average marginal reliability results:
Biological and Physical Sciences

<table>
<thead>
<tr>
<th>Source</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulated for 63 linear items</td>
<td>0.85</td>
</tr>
<tr>
<td>Simulated for 48 linear items</td>
<td>0.82</td>
</tr>
<tr>
<td>Simulated for 48 3 stage 3 level items</td>
<td>0.86</td>
</tr>
<tr>
<td>Simulated for 48 3 stage 4 level items</td>
<td>0.86</td>
</tr>
</tbody>
</table>

The Target Information Function (TIF)

General Findings and Final Thoughts

- General Results
  - Bio/Phys: 15 item reduction (63 to 48)
  - Verbal: 14 item reduction (50 to 36)
- Item Discrimination and CAT
- Reduction in overall length
  - Psychometric implications
  - Test taker implications
- Uniform item exposure
  - No "on-the-fly" solution required
  - Balanced use of items in pool