What is an item bank?

A database?
- physical store of test questions
- computerised store of test questions
- information about items (e.g., topic, date written, name of author, date last used, item statistics etc)

A set of items measuring the same trait which have been calibrated onto the same scale
- make tests of different content / difficulty
- different raw scores → same ability estimate

Calibrated tests

Demonstration using RUMM
Adding items to the bank

Data collection design is important
MUST have a ‘link’ in order to bring new items onto
the same scale

Common person link
• the same people take both new and old items
• ‘strong’ link easier to obtain by increasing N

Common item link
• some items in both new and old test forms
• more practical - produce tests of the same length

Common person link

<table>
<thead>
<tr>
<th>Persons</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 2 3 4 5 ... ... ... ... L</td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td></td>
</tr>
</tbody>
</table>

Common item link

<table>
<thead>
<tr>
<th>Persons</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 2 3 4 5 ... ... ... ... L</td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td></td>
</tr>
</tbody>
</table>
Complex linking design

<table>
<thead>
<tr>
<th>Persons</th>
<th>1 2 3 4 5 ...</th>
<th>...</th>
<th>...</th>
<th>...</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3 linking methods

Linking with anchor items
Linking with ‘1-step equating’
  - ‘Joint calibration’ or ‘concurrent estimation’
Linking with ‘2-step equating’
  - ‘Separate calibration’

Linking with anchor items
Fix some items at a known calibration

Advantages:
  - quick and flexible
  - best with a large, well-established item bank

Disadvantages:
  - estimation algorithm might not converge
  - might have to run analysis several times
  - less control for the analyst
Linking with 1-step equating

Arrange data from linked tests into a single large matrix with missing data

Advantages:
• quick - estimation algorithm handles missing data
• resulting calibrations can be banked immediately

Disadvantages:
• Unit size differences compared to 2-step
• Problems if equating 'vertically' i.e. items cover a wide range of difficulty
• Can be difficult to prepare data set if linking items appear in different positions on different test forms

Linking with 2-step equating

Analyse tests separately then calculate adjustment needed to align scales

Advantages
• Control and understanding for the analyst
• Easy to identify link items which do not maintain their calibration across test forms

Disadvantages
• Time-consuming, possibility of human error
• Need to calculate single calibration for common items before banking them
• Not practical if number of test forms > 3

How to link two tests (1)

1. Choose which test will form the basis e.g. put Test B onto same scale as Test A
2. Identify the common items
3. Calculate the mean difficulty of the common items in each test, $C_A$ and $C_B$
4. Calculate the 'link constant', $L = C_A - C_B$
5. Add the link constant $L$ to ALL item difficulties in Test B
6. Plot the new Test B common item difficulties against the Test A common item difficulties
How to link two tests (2)

7. Add the line $y=x$ to the plot
8. Add some ‘error tramlines’ to the plot
9. Identify any poor linking items
10. Re-calculate $L$ without these items
11. Repeat steps 6-10 until you are satisfied
12. Calculate new average difficulty for common items
13. Test A and Test B items now form a calibrated item bank!

Example – steps 1 & 2

<table>
<thead>
<tr>
<th>Test A</th>
<th>Difficulty</th>
<th>Test B</th>
<th>Difficulty</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>-2.3</td>
<td>B1</td>
<td>-2.2</td>
</tr>
<tr>
<td>A2</td>
<td>-3.1</td>
<td>B2</td>
<td>-2.7</td>
</tr>
<tr>
<td>A3</td>
<td>-1.8</td>
<td>B3</td>
<td>-2.1</td>
</tr>
<tr>
<td>A4</td>
<td>-0.3</td>
<td>B4</td>
<td>-1</td>
</tr>
<tr>
<td>A5</td>
<td>-1.1</td>
<td>B5</td>
<td>-0.3</td>
</tr>
<tr>
<td>A6</td>
<td>0.5</td>
<td>B6</td>
<td>0.8</td>
</tr>
<tr>
<td>A7</td>
<td>1.2</td>
<td>B7</td>
<td>1.1</td>
</tr>
<tr>
<td>A8</td>
<td>1.4</td>
<td>B8</td>
<td>1.6</td>
</tr>
<tr>
<td>A9</td>
<td>2.2</td>
<td>B9</td>
<td>2.9</td>
</tr>
<tr>
<td>A10</td>
<td>3.3</td>
<td>B10</td>
<td>1.9</td>
</tr>
<tr>
<td>Mean</td>
<td>0.0</td>
<td>Mean</td>
<td>0.0</td>
</tr>
</tbody>
</table>