Item Response Theory and Computerized Adaptive Testing

Richard C. Gershon, PhD
Department of Medical Social Sciences
Feinberg School of Medicine
Northwestern University
gershon@northwestern.edu
May 20, 2011
Outline

♦ Item Response Theory
  ♦ versus Classical Test Theory

♦ Uses of IRT
  ♦ Item Banking
  ♦ Short Forms
  ♦ Computerized Adaptive Tests
Psychometrics is the field of study concerned with the theory and technique of educational and psychological measurement, which includes the measurement of knowledge (achievement), abilities, attitudes, and personality traits.

Measurement requires the concept of an underlying trait that can be expressed in terms of more or less. Test items are the operational definition of the underlying trait. Test items can be ordered from easy to hard. Test takers can be ordered from less able to more able.
### IRT Modeling is Latent Trait Modeling

- A latent trait is an *unobservable* latent dimension that is thought to give rise to a set of observed item responses.

#### I am too tired to do errands

<table>
<thead>
<tr>
<th></th>
<th>False</th>
<th>True</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Fatigue**
These latent traits (constructs, variables, $\theta$) are measured on a continuum of severity.

I am too tired to do errands?

energetic  False  True  severe

Fatigue
Advantages of Using IRT

- Equal Interval Measure
- Test-takers and items are represented on the same scale
- Item calibrations are independent of the test-takers used for calibration
- Candidate ability estimates are independent of the particular set of items used for estimation
- Measurement precision is estimated for each person and each item
Test-takers and Items are Represented on the Same Scale

- Item Difficulty = Severity = Measure = Theta = Item Calibration = Location

- Person Ability = Measure = Theta = Person Calibration = Location
Test-takers and Items are Represented on the Same Scale

**Item Difficulty**

I have trouble getting out of bed.  
I have too tired to do errands.  
I can easily run marathons.

<table>
<thead>
<tr>
<th>Easy</th>
<th>Low</th>
<th>Person QOL</th>
<th>Hard</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- arrows pointing down: Easy
- arrows pointing up: High
More Basic Terms

- Discrimination = the degree to which an item discriminates person ability
- Item Information = the area where an item discriminates
- Test Information = the area where the test discriminates
Last, but not Least

- Item “Parameters” =
  - IRT statistics about an item
  - Primary: Item Difficulty
  - Often: Item Discrimination
  - Sometimes: Guessing
  - Lots of other “ugly looking numbers”
Differential Item Functioning (DIF)

- Does an item have different item parameters for different subgroups?
  - Gender
  - Race
  - Age
  - Disease
The Three Main IRT Models

- Rasch model one parameter logistic model (1PL)
- Two parameter logistic model (2PL)
- Three parameter logistic model (3PL)
How to choose an appropriate IRT Model

OR

My religion is better than your religion!
WARNING!

- You are about to see mathematical formulas!
When the difficulty of a given item exactly matches the Examinee’s ability level, then the person has 50% chance of answering that item correctly:
One Parameter Logistic Model

- Only option for small sample sizes
- Often the real model underlying a test labeled as three parameter
- Less costly
- “The simple solution is always the best”
Two Parameter Logistic Model

\[ P_{1,0} = \frac{e^{a(\text{ability} - b)}}{1 + e^{a(\text{ability} - b)}} \]

Two parameters
- \( a = \text{Discrimination} \)
- \( b = \text{Item Difficulty} \)
Two Parameter Examples

- $a=0.5, b=0.5, c=0.1$
- $a=1.5, b=0.5, c=0.1$
- $a=2.5, b=0.5, c=0.1$
Three Parameter Logistic Model

\[ P_{1,0} = c + (1-c) \frac{e^{a (ability - b)}}{1 + e^{a (ability - b)}} \]

Three parameters
- \( a \): Discrimination
- \( b \): Item Difficulty
- \( c \): Guessing
Three Parameter Logistic Model (3PL)

- Requires a large sample size
- Significant research demonstrating that theoretically 3PL is better, but practically has little advantage over 1PL
- “Most accepted theoretical model”
Three Parameter Examples

\[ a=1.5, b=0.5, c=0.1 \]

\[ a=2.5, b=0.5, c=0.25 \]
Polytomous Models

One Parameter
- Rating Scale Model
- Partial Credit Model

Two Parameter
- Graded Response Model
- Generalized Partial Credit Model
Multi-dimensional Models

- There are also IRT models which consider more than one unidimensional trait at a time
How does IRT differ from conventional test theory?
An individual takes an assessment
Their total score on that assessment is used for comparison purposes
High Score – The person is higher on the trait
Low Score - The person is lower on the trait
Item Response Theory

- Each individual item can be used for comparison purposes
- Person endorses better rating on “hard items” - The person is higher on the trait
- Person endorses worse rating on “easy items” - The person is lower on the trait
- Items that measure the same construct can be aggregated into longer assessments
# Reliability

### CTT
- Reliability is based upon the total test.
- Regardless of patient “ability”, reliability is the same.

### IRT
- Reliability is calculated for each patient “ability” and varies across the continuum.
- Typically, there is better reliability in the middle of the distribution.
# Validity

<table>
<thead>
<tr>
<th>CTT</th>
<th>IRT</th>
</tr>
</thead>
<tbody>
<tr>
<td>✷ Validity is based upon the total test.</td>
<td></td>
</tr>
<tr>
<td>✷ Typically, validity would need to be re-assessed if the instrument is modified in any way.</td>
<td></td>
</tr>
<tr>
<td>✷ Validity is assessed for the entire item bank.</td>
<td></td>
</tr>
<tr>
<td>✷ Subsets of items (full length tests, short forms and CAT) <strong>all</strong> inherit the validity assessed for the original item bank.</td>
<td></td>
</tr>
</tbody>
</table>
How Scores Depend on the Difficulty of Test Items

Very Easy Test

Person

Expected Score 8

1 8

1 8

Expected Score 0

Very Hard Test

Person

Medium Test

Person

Expected Score 5

1 8

Raw Scores vs. IRT Measures
IRT has Equal Interval Measurement

4 Item Test

Raw:

Logit Measures:
How Differences Between Person Ability and Item Difficulty Ought to Affect the Probability of a Correct Response

The Item Characteristic Curve

Sample Item Characteristic Curve (ICC)
I Have a Lack of Energy

Traditional Test Theory

0 = Very Much  1 = Quite a Bit  2 = Somewhat  3 = A Little Bit  4 = Not at All
I Have a Lack of Energy

Traditional Test Theory

0 = Very Much  1 = Quite a Bit  2 = Somewhat  3 = A Little Bit  4 = Not at All

Item Response Theory
The IRT “Reality” of a 10 Point Rating-Scale Item

No Pain

Worst Pain
I have a lack of energy

This is an Item Characteristic Curve (ICC) for a rating scale item (each option has its own curve)

0 = Very Much  1 = Quite a Bit  2 = Somewhat  3 = A Little Bit  4 = Not at All
I have a lack of energy

0 = Very Much; 1 = Quite a Bit; 2 = Somewhat; 3 = A Little Bit; 4 = Not at All
I have a lack of energy

0 = Very Much; 1 = Quite a Bit; 2 = Somewhat; 3 = A Little Bit; 4 = Not at All
I have a lack of energy

0 = Very Much; 1 = Quite a Bit; 2 = Somewhat; 3 = A Little Bit; 4 = Not at All
I have a lack of energy

0 = Very Much; 1 = Quite a Bit; 2 = Somewhat; 3 = A Little Bit; 4 = Not at All
0 = Very Much; 1 = Quite a Bit; 2 = Somewhat; 3 = A Little Bit; 4 = Not at All
IRT Polytomous Responses

I have been too tired to feel happy.

Fatigue

<table>
<thead>
<tr>
<th>Probability of Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>None of the time</td>
</tr>
<tr>
<td>A little of the time</td>
</tr>
<tr>
<td>Some of the time</td>
</tr>
<tr>
<td>Most of the time</td>
</tr>
<tr>
<td>All of the time</td>
</tr>
</tbody>
</table>

Fatigue

<table>
<thead>
<tr>
<th>Energetic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe Fatigue</td>
</tr>
</tbody>
</table>
IRT Polytomous Responses
I have felt energetic.

Fatigue
-3.00 -2.00 -1.00 0.00 1.00 2.00 3.00
Fatigue

Most of the time
Energetic
All of the time
Some of the time
A little of the time
None of the time

Probability of Response

Fatigue

Energetic
Severe Fatigue
IRT Polytomous Responses

I have been too tired to read.

Fatigue

Probability of Response

None of the time
A little of the time
Some of the time
Most of the time
All of the time

Energetic
Severe Fatigue

Probability of Response

0.0
0.2
0.4
0.6
0.8
1.0

-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00

Fatigue

0

Fatigue

0
ITEM BANKING
Calibrated Item Banks can be used to Create Numerous Instrument Types

Short Forms
- 5-7 Items in each HRQL Area
- Constructed to cover full range of trait
  OR
- Multiple forms constructed to only cover a narrow range of trait (e.g., high, medium, or low)

Emotional Distress
- Physical Function
  Item Bank
  - Item40
  - Item38
  - Item36
  - Item34
  - Item32
  - Item30
  - Item28
  - Item26
  - Item24
  - Item22
  - Item20
  - Item18
  - Item16
  - Item14
  - Item12
  - Item10
  - Item8
  - Item6
  - Item4
  - Item2

Pain

Computerized Adaptive Testing (CAT)
- Custom individualized assessment
- Suitable for clinical use
- Accuracy level chosen by researcher

Custom Item Selection

Prostate Cancer
- Item40
- Item38
- Item34
- Item32
- Item26
- Item22
- Item18
- Item16
- Item14
- Item12
- Item10
- Item8
- Item6
- Item4
- Item2

Breast Cancer
- Item36
- Item34
- Item32
- Item28
- Item26
- Item22
- Item18
- Item16
- Item14
- Item10
- Item8
- Item6
- Item4
- Item2

Brain Tumor
- Item40
- Item32
- Item24
- Item22
- Item16
- Item8
- Item2

- 3 Diseases
- 3 Trials
- 3 Unique Instruments
  - Each based on content interest of individual researchers

5-7 Items in each HRQL Area

- Constructed to cover full range of trait
- OR
- Multiple forms constructed to only cover a narrow range of trait (e.g., high, medium, or low)

Depression

Depressive Symptoms Item Bank

no depression mild depression moderate depression severe depression extreme depression
Computerized Adaptive Testing (CAT)

- Emotional Distress
- Pain
- Physical Function

Item Bank
- Item 40
- Item 38
- Item 36
- Item 34
- Item 32
- Item 30
- Item 28
- Item 26
- Item 24
- Item 22
- Item 20
- Item 18
- Item 16
- Item 14
- Item 12
- Item 10
- Item 8
- Item 6
- Item 4
- Item 2

- Custom individualized assessment
- Suitable for clinical use
- Accuracy level chosen by researcher

# Custom Item Selection

### Emotional Distress
- Pain
- Physical Function
  - Item Bank
    - Item 40
    - Item 38
    - Item 36
    - Item 34
    - Item 32
    - Item 30
    - Item 28
    - Item 26
    - Item 24
    - Item 22
    - Item 20
    - Item 18
    - Item 16
    - Item 14
    - Item 12
    - Item 10
    - Item 8
    - Item 6
    - Item 4
    - Item 2

### Custom Item Selection

<table>
<thead>
<tr>
<th>Prostate Cancer</th>
<th>Breast Cancer</th>
<th>Brain Tumor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 40</td>
<td>Item 36</td>
<td>Item 40</td>
</tr>
<tr>
<td>Item 38</td>
<td>Item 34</td>
<td>Item 32</td>
</tr>
<tr>
<td>Item 34</td>
<td>Item 34</td>
<td>Item 32</td>
</tr>
<tr>
<td>Item 32</td>
<td>Item 32</td>
<td>Item 28</td>
</tr>
<tr>
<td>Item 26</td>
<td>Item 26</td>
<td>Item 26</td>
</tr>
<tr>
<td>Item 22</td>
<td>Item 22</td>
<td>Item 24</td>
</tr>
<tr>
<td>Item 18</td>
<td>Item 18</td>
<td>Item 16</td>
</tr>
<tr>
<td>Item 16</td>
<td>Item 16</td>
<td>Item 14</td>
</tr>
<tr>
<td>Item 8</td>
<td>Item 8</td>
<td>Item 10</td>
</tr>
<tr>
<td>Item 2</td>
<td>Item 2</td>
<td>Item 8</td>
</tr>
</tbody>
</table>

- **3 Diseases**
- **3 Trials**
- **3 Unique Instruments**
  - Each based on content interest of individual researchers

In Summary,
Calibrated Item Banks can be used to:

- Create a standard static instrument
- Construct short forms
- Enable CAT
- Select items based on unique content interests and formulate custom short-form or full-length instruments
In every case, using a validated, pre-calibrated item bank allows any of these instruments to be pre-validated and produce standardized scores on the same scale.
Computerized Adaptive Testing
What is Computerized Adaptive Testing?

- Shorter
- Targeting
- Computerized Algorithm
CAT in the Military
Armed Services Vocational Aptitude Battery (ASVAB)
CAT for Certification

American Association of Nurse Anesthetists

Microsoft

CompTIA

Society of Clinical Pharmacology and Clinician's Research
CAT for Licensure

AMERICAN DIETETIC ASSOCIATION

NASD

National Council of State Boards of Nursing, Inc.
CAT for College Entrance

The GMAT

The College Board

ACCUPLACER OnLine

toefl.org

Test of English as a Foreign Language

Graduate Record Examinations®
CAT for Education

Northwest Evaluation Association

Renaissance Learning™
CAT for Clinical Testing
CAT for Personnel Testing
Example – Binary Search

- Binary search
Result:
Medium-High on the Trait
When is CAT appropriate?

- Heterogeneous populations
- Diagnostic tests
- On-demand testing
- Long tests
When is CAT not appropriate?

- Small populations
  - Difficulty in calibrating items
  - Higher administration costs
CAT Requirements

- Calibrated item bank
- Administration software
How do I create a calibrated item bank?

- You probably already have done the hard work!
- It’s usually the same as for CBT.
Create an Item Bank
Part 1

- Item sources
  - Previous exams
    - With a CAT bank fewer constraints exist regarding item re-use
  - Write new items
Create an Item Bank
Part 2

- Item quality
  - Statistics relevant to CAT (not necessarily print)
    - Difficulty
    - Other variables relevant to selected IRT model
Calibrating items
Part 1

- Analyze item level data from previously administered items
  - Preferably using raw person-data
  - Alternatively, could create raw estimates from p-values
  - Several software packages exist for this purpose
Calibrating items

Part 2

- Pilot (beta) new test items
  - On paper
  - On computer

- Typically requires a psychometrician to supervise analyses
Test Specifications

- Starting rule
  - With item which provides maximum information
  - At cut point
Test Specifications

- Stopping Rule
  - Fixed length
  - Variable length
  - by Total Test/Subtest
  - Calculated
    - Specified precision of measure
    - Specified confidence in a pass/fail decision
- Maximum item count
- Minimum item count
Test Specifications

- Content balancing
  - None
  - Fixed percentage
Test Specifications

- Testing new items (beta testing, field testing, experimental items)
Adaptive Algorithm

- Person ability algorithm
- Item selection algorithm
  - Test difficulty
  - Maximum jump size
  - Content issues
  - Item exposure control
  - Option to not allow same items to be used during retesting
  - Overlapping items (items that cue other items)
<table>
<thead>
<tr>
<th>Item</th>
<th>AN Cont</th>
<th>Diff Ans</th>
<th>Time</th>
<th>Meas</th>
<th>SE</th>
<th>0</th>
<th>S</th>
<th>1.0</th>
<th>2.0</th>
<th>3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31384</td>
<td>ST</td>
<td>0.35</td>
<td>1</td>
<td>0</td>
<td>0'34</td>
<td>-9.99</td>
<td>9.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>31009</td>
<td>FIX</td>
<td>0.22</td>
<td>3</td>
<td>1</td>
<td>0'18</td>
<td>0.29</td>
<td>1.41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>31113</td>
<td>LO</td>
<td>0.18</td>
<td>1</td>
<td>0</td>
<td>0'26</td>
<td>-0.44</td>
<td>1.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>30385</td>
<td>MIC</td>
<td>0.36</td>
<td>3</td>
<td>1</td>
<td>0'33</td>
<td>0.28</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>30873</td>
<td>ST</td>
<td>0.24</td>
<td>3</td>
<td>0</td>
<td>0'31</td>
<td>-0.14</td>
<td>0.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>30533</td>
<td>PRO</td>
<td>0.05</td>
<td>2</td>
<td>0</td>
<td>0'30</td>
<td>-0.46</td>
<td>0.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>30525</td>
<td>ST</td>
<td>0.35</td>
<td>2</td>
<td>0</td>
<td>0'16</td>
<td>-0.67</td>
<td>0.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>31008</td>
<td>FIX</td>
<td>0.37</td>
<td>4</td>
<td>0</td>
<td>0'31</td>
<td>-0.83</td>
<td>0.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>30664</td>
<td>ST</td>
<td>0.30</td>
<td>2</td>
<td>0</td>
<td>0'32</td>
<td>-0.98</td>
<td>0.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>31086</td>
<td>LO</td>
<td>0.35</td>
<td>4</td>
<td>0</td>
<td>0'12</td>
<td>-1.11</td>
<td>0.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>31626</td>
<td>ST</td>
<td>0.34</td>
<td>2</td>
<td>0</td>
<td>0'23</td>
<td>-1.22</td>
<td>0.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>31356</td>
<td>MIC</td>
<td>0.32</td>
<td>4</td>
<td>1</td>
<td>0'41</td>
<td>-0.81</td>
<td>0.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>31210</td>
<td>PRO</td>
<td>0.21</td>
<td>2</td>
<td>0</td>
<td>0'35</td>
<td>-0.92</td>
<td>0.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>31148</td>
<td>ST</td>
<td>0.39</td>
<td>1</td>
<td>0</td>
<td>0'20</td>
<td>-1.01</td>
<td>0.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>31620</td>
<td>FIX</td>
<td>0.25</td>
<td>4</td>
<td>0</td>
<td>0'10</td>
<td>-1.10</td>
<td>0.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>30224</td>
<td>ST</td>
<td>0.20</td>
<td>4</td>
<td>0</td>
<td>0'25</td>
<td>-1.19</td>
<td>0.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>30940</td>
<td>FIX</td>
<td>0.40</td>
<td>2</td>
<td>0</td>
<td>0'32</td>
<td>-1.25</td>
<td>0.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>31288</td>
<td>ST</td>
<td>0.25</td>
<td>3</td>
<td>1</td>
<td>1'14</td>
<td>-0.97</td>
<td>0.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>31529</td>
<td>LO</td>
<td>0.28</td>
<td>1</td>
<td>0</td>
<td>0'58</td>
<td>-1.04</td>
<td>0.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>31120</td>
<td>ST</td>
<td>0.40</td>
<td>2</td>
<td>0</td>
<td>0'11</td>
<td>-1.10</td>
<td>0.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>31355</td>
<td>MIC</td>
<td>0.36</td>
<td>2</td>
<td>0</td>
<td>0'59</td>
<td>-1.15</td>
<td>0.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>31207</td>
<td>PRO</td>
<td>0.33</td>
<td>2</td>
<td>0</td>
<td>0'34</td>
<td>-1.21</td>
<td>0.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>30745</td>
<td>ST</td>
<td>0.33</td>
<td>4</td>
<td>0</td>
<td>0'33</td>
<td>-1.26</td>
<td>0.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>31285</td>
<td>FIX</td>
<td>0.40</td>
<td>3</td>
<td>0</td>
<td>0'13</td>
<td>-1.31</td>
<td>0.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>30237</td>
<td>ST</td>
<td>0.39</td>
<td>3</td>
<td>0</td>
<td>0'22</td>
<td>-1.35</td>
<td>0.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>30179</td>
<td>ST</td>
<td>0.26</td>
<td>1</td>
<td>0</td>
<td>0'24</td>
<td>-1.40</td>
<td>0.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>31055</td>
<td>FIX</td>
<td>0.23</td>
<td>4</td>
<td>1</td>
<td>0'24</td>
<td>-1.18</td>
<td>0.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>31598</td>
<td>LO</td>
<td>0.29</td>
<td>2</td>
<td>0</td>
<td>0'33</td>
<td>-1.23</td>
<td>0.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>30384</td>
<td>MIC</td>
<td>0.27</td>
<td>3</td>
<td>0</td>
<td>0'11</td>
<td>-1.27</td>
<td>0.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>30524</td>
<td>ST</td>
<td>0.20</td>
<td>1</td>
<td>0</td>
<td>0'21</td>
<td>-1.31</td>
<td>0.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>31470</td>
<td>PRO</td>
<td>0.38</td>
<td>1</td>
<td>0</td>
<td>0'16</td>
<td>-1.35</td>
<td>0.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>30188</td>
<td>ST</td>
<td>0.31</td>
<td>3</td>
<td>0</td>
<td>0'21</td>
<td>-1.39</td>
<td>0.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>31402</td>
<td>FIX</td>
<td>0.28</td>
<td>3</td>
<td>0</td>
<td>1' 9</td>
<td>-1.42</td>
<td>0.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>AN</td>
<td>Cont</td>
<td>Diff Ans</td>
<td>Time</td>
<td>Meas</td>
<td>SE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>------</td>
<td>----------</td>
<td>------</td>
<td>------</td>
<td>----</td>
<td>-----</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>220576 SC</td>
<td>-0.33</td>
<td>3 1 o</td>
<td>0'37</td>
<td>9.99</td>
<td>9.99</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>220304 LO</td>
<td>-0.24</td>
<td>1 1 o</td>
<td>1'13</td>
<td>9.99</td>
<td>9.99</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>220935 SPH</td>
<td>-0.13</td>
<td>4 0 o</td>
<td>1'3</td>
<td>0.46</td>
<td>1.22</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>220213 SC</td>
<td>-0.03</td>
<td>1 1 +</td>
<td>0'52</td>
<td>0.92</td>
<td>1.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>220378 AP</td>
<td>-0.11</td>
<td>3 0 =</td>
<td>0'40</td>
<td>0.24</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>220523 SC</td>
<td>-0.30</td>
<td>4 1 o</td>
<td>0'10</td>
<td>0.50</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>220611 LO</td>
<td>-0.37</td>
<td>1 1 o</td>
<td>1'17</td>
<td>0.70</td>
<td>0.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>220928 SC</td>
<td>-0.38</td>
<td>1 0 o</td>
<td>0'33</td>
<td>0.27</td>
<td>0.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>220218 SPH</td>
<td>-0.48</td>
<td>3 0 o</td>
<td>0'50</td>
<td>-0.04</td>
<td>0.67</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>220975 SC</td>
<td>-0.65</td>
<td>3 1 o</td>
<td>0'46</td>
<td>0.10</td>
<td>0.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>220709 SC</td>
<td>-0.79</td>
<td>1 1 o</td>
<td>0'35</td>
<td>0.21</td>
<td>0.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>220634 LO</td>
<td>-0.56</td>
<td>2 0 =</td>
<td>0'41</td>
<td>-0.03</td>
<td>0.59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>220708 SPH</td>
<td>-0.81</td>
<td>1 1 o</td>
<td>0'22</td>
<td>0.07</td>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>220748 SC</td>
<td>-0.65</td>
<td>2 0 o</td>
<td>0'34</td>
<td>-0.13</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>220369 AP</td>
<td>-0.88</td>
<td>2 1 o</td>
<td>0'39</td>
<td>-0.04</td>
<td>0.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>220777 SC</td>
<td>-0.68</td>
<td>1 0 o</td>
<td>0'40</td>
<td>-0.21</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>220265 LO</td>
<td>-0.97</td>
<td>1 0 o</td>
<td>0'12</td>
<td>-0.37</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>220885 SC</td>
<td>-0.95</td>
<td>1 1 o</td>
<td>0'33</td>
<td>-0.29</td>
<td>0.47</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>220302 SPH</td>
<td>-0.98</td>
<td>2 1 o</td>
<td>0' 8</td>
<td>-0.22</td>
<td>0.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>220044 SC</td>
<td>-0.88</td>
<td>1 1 o</td>
<td>0'32</td>
<td>-0.15</td>
<td>0.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>220442 SC</td>
<td>-0.80</td>
<td>4 0 o</td>
<td>0'16</td>
<td>-0.28</td>
<td>0.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>220263 LO</td>
<td>-1.01</td>
<td>1 1 o</td>
<td>0'52</td>
<td>-0.22</td>
<td>0.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>220507 SPH</td>
<td>-0.79</td>
<td>1 0 o</td>
<td>0'30</td>
<td>-0.34</td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>220037 SC</td>
<td>-1.00</td>
<td>4 1 +</td>
<td>0'43</td>
<td>-0.28</td>
<td>0.41</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>220317 AP</td>
<td>-1.05</td>
<td>3 1 o</td>
<td>0'11</td>
<td>-0.23</td>
<td>0.41</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>220535 SC</td>
<td>-0.92</td>
<td>3 0 =</td>
<td>0'51</td>
<td>-0.33</td>
<td>0.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>220987 LO</td>
<td>-1.02</td>
<td>4 1 o</td>
<td>0'25</td>
<td>-0.28</td>
<td>0.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>220342 SC</td>
<td>-0.99</td>
<td>3 1 o</td>
<td>0'49</td>
<td>-0.23</td>
<td>0.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>220089 SPH</td>
<td>-0.89</td>
<td>2 0 o</td>
<td>0'41</td>
<td>-0.33</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>220860 SC</td>
<td>-1.11</td>
<td>2 1 o</td>
<td>0'20</td>
<td>-0.29</td>
<td>0.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>220754 SC</td>
<td>-0.98</td>
<td>3 0 o</td>
<td>0'47</td>
<td>-0.38</td>
<td>0.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>220610 LO</td>
<td>-1.08</td>
<td>3 1 o</td>
<td>0'23</td>
<td>-0.33</td>
<td>0.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>220347 SPH</td>
<td>-0.91</td>
<td>1 1 o</td>
<td>0'49</td>
<td>-0.29</td>
<td>0.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>220856 SC</td>
<td>-1.01</td>
<td>2 1 +</td>
<td>1' 2</td>
<td>-0.25</td>
<td>0.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
GP1 – I have a lack of energy

0 = Very Much; 1 = Quite a Bit; 2 = Somewhat; 3 = A Little Bit; 4 = Not at All
GP1 – I have a lack of energy

0 = Very Much; 1 = Quite a Bit; 2 = Somewhat; 3 = A Little Bit; 4 = Not at All
Simulate Measure = 48

Information Function

<table>
<thead>
<tr>
<th>Item</th>
<th>Meas</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>37</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Simulate Measure = 48
Simulate Measure = 48

Information Function

<table>
<thead>
<tr>
<th>Item</th>
<th>Meas</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>42</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>Meas</td>
<td>SE</td>
</tr>
<tr>
<td>------</td>
<td>------</td>
<td>-----</td>
</tr>
<tr>
<td>1</td>
<td>37</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>42</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>44</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>45</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>46</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>47</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>48</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>47</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>48</td>
<td>5</td>
</tr>
<tr>
<td>Item</td>
<td>Meas</td>
<td>SE</td>
</tr>
<tr>
<td>------</td>
<td>------</td>
<td>----</td>
</tr>
<tr>
<td>1</td>
<td>9</td>
<td>38</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>19</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>16</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>17</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>15</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>14</td>
<td>5</td>
</tr>
</tbody>
</table>

Simulate Measure = 15
Simulate Measure = 92

<table>
<thead>
<tr>
<th>Item</th>
<th>Meas</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>25</td>
</tr>
<tr>
<td>2</td>
<td>97</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>99</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>97</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>95</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>97</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>99</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>100</td>
<td>7</td>
</tr>
<tr>
<td>10</td>
<td>99</td>
<td>7</td>
</tr>
</tbody>
</table>

This is called a “ceiling” effect. The test doesn’t have good items to assess people at the “ceiling” of the trait range.

If the test didn’t have items to assess people at the bottom of the trait range this would be called a “floor” effect.
### Sample PROMIS Fatigue Short Form

<table>
<thead>
<tr>
<th>Question</th>
<th>Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Often</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>How often did you feel tired?</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>How often did you experience extreme exhaustion?</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>How often did you run out of energy?</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>How often did your fatigue limit you at work (include work at home)?</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>How often were you too tired to think clearly?</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>How often were you too tired to take a bath or shower?</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>How often did you have enough energy to exercise strenuously?</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
</tbody>
</table>

In the past 7 days ...

Reprinted with permission of the PROMIS Health Organization and the PROMIS Cooperative Group © 2007.
Demonstration

- CAT in Assessment Center