Latent variables

- The Theory of True Scores
  - Sometimes called or ‘Latent Trait Theory’
- \( X = T + E \)
  - Where \( X \) = Observed Score
  - \( T \) = True Score
  - \( E \) = Error
- Latent Variable Analysis (Factor Analysis)
The Psychometric Principles
Maximizing the quality of assessment

• Reliability (freedom from error)
• Validity (‘... what is says on the tin’)
• Standardisation (compared with what?)
• Equivalence (is it biased?)

Constructing a psychometric test

- Defining the purpose
- Designing the blueprint
- The pilot study
- Item analysis
- Obtaining reliability and validity
- Writing the handbook
Appropriate latent traits

- Personality rather than ability
  - e.g. “Tendency to feel ….”
- Single traits
- About 32 pilot items
- Reduce to 20 through item analysis
- Some ideas
Openness

Openness-to-Experience

I am always on the lookout for new ideas to explore

Conventionalism

Conventionality

Stick with tried and tested ways of doing things
Conscientious

Detail conscious
If you get the details right everything will come out fine.

Seeing the bigger picture
I prefer to leave boring details to others

Disorganised
Extraversion

I can be the life and soul of any party

Introversion

I have my best ideas when I am left alone
Agreeableness

Tendermindedness

I will always try and help my friends when they need me

Disagreeable

I have a special ability to make tough decisions when needed

Toughmindedness
Neurotic

Emotional
I get anxious when I have to make difficult decisions

Unemotional
I am more able than most to cope with disaster

Stable
Prudence

Proficiency
Careful and sensible when carrying out given tasks

Carelessness
May be careless in dealing with mundane tasks

Foolishness
Fortitude

Work-orientation
An effective and confident orientation to hard work

Absenteeism
Can dislike working under pressure

Inconstancy
Temperance

Anger

**Patience**

Slow to anger and calm in the face of adversity

**Hostility**

Sometimes impatience with the foibles of others
Justice

Fair-mindedness
Balanced and impartial in decision making

Subversion
Tends to be suspicious of the intentions of others

Envy
Loyalty
Shows obedience to others and a sense of duty

Disloyalty
Might ignore the advice of those who know better
Charity

Generosity
Honest and open in dealings with other

Covetousness
Occasionally prone to elaborate or distort the truth

Greed
Hope

Resourceful and enthusiastic about the future

Desolation

Inertia

Somewhat pessimistic about the future
The book on human emotions
*Tiffany Watt Smith, Profile Books: London, 2015*

- Apathy,
- Bewilderment
- Collywobbles
- Disgruntlement
- Embarrassment
- feeling like a Fraud
- Guilt
- Hatred
- Humility
- Indignation
- Jealousy
- Loneliness
- Love
- Morbid Curiosity
- Nostalgia
- Overwhelmed
- Perversity
- Road Rage
- Self-Pity
- Triumph
- Uncertainty
- Vengefulness
- Wanderlust
Developing the blueprint

• Curriculum based
  • Bloom’s taxonomy of educational objectives

• Job description
  • The job analysis
  • The person specification

• Theoretical
  • Ability
  • Personality
# Knowledge Test Specification

## Content areas

<table>
<thead>
<tr>
<th>Manifestations</th>
<th>Arithmetic</th>
<th>Geometry</th>
<th>Algebra</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge of Terms (25%)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Understanding (25%)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Application (25%)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Generalisation (25%)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>
## Personality Test Specification

<table>
<thead>
<tr>
<th>Manifestations</th>
<th>Extraversion</th>
<th>Neuroticism</th>
<th>Detail</th>
<th>Tough-mindedness</th>
</tr>
</thead>
<tbody>
<tr>
<td>High/Positive</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>High/Negative</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Low/Positive</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Low/Negative</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>
The Pilot Study

- Pre-piloting
  - Are the correct items correct?
  - Are the distractors incorrect?
  - Are any items offensive or likely to be biased?
- The sample and sample size
- Data collection
- Data entry
<table>
<thead>
<tr>
<th>Classical</th>
<th>Item Response Theory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difficulty Value ( (p) )</td>
<td>Rasch Parameter ( (1^{st}) )</td>
</tr>
<tr>
<td>Discrimination Index ( (r) )</td>
<td>Discrimination Parameter ( (2^{nd}) )</td>
</tr>
<tr>
<td>Correction for Guessing</td>
<td>Guessing Parameter ( (3^{rd}) )</td>
</tr>
<tr>
<td></td>
<td>Inattention Parameter ( (4^{th}) )</td>
</tr>
</tbody>
</table>
Using difficulty \((p)\) and discrimination \((r)\) Indices

- \(p\) should be between 20% and 80%
- Adjusted Item-Total Correlation \(> 0.2\)
- Remember the test specification!
Classical item reduction

- Record form analysis
  - Non-responses
  - Altered items
  - Comments
- Delete extreme items
- Delete items with poor discrimination
- Retain the balance of the test
  - Test specification
  - Positive and negative items
- Aim to reduce items by 50%
Reliability

- Inter-rater reliability
- Test–retest reliability (stability)
- Parallel forms reliability
- Split-half reliability
  - The Spearman-Brown Formula
- Cronbach’s alpha
## Expected levels of reliability

<table>
<thead>
<tr>
<th>Test Type</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>WAIS</td>
<td>0.92</td>
</tr>
<tr>
<td>Ability tests</td>
<td>0.82</td>
</tr>
<tr>
<td>Personality test</td>
<td>0.75</td>
</tr>
<tr>
<td>Essay marks</td>
<td>0.60</td>
</tr>
<tr>
<td>Supervisors ratings</td>
<td>0.40</td>
</tr>
<tr>
<td>Projective tests</td>
<td>0.30</td>
</tr>
<tr>
<td>Creativity tests</td>
<td>0.25</td>
</tr>
<tr>
<td>Astrology/graphology?</td>
<td>?</td>
</tr>
</tbody>
</table>
• p should be between 20% and 80%
• Adjusted Item-Total Correlation (point or rank biserial) $> 0.2$
• Remember the test specification!
Software

- R
- Excel (with Real Statistic or other Add On)
- Mplus
- SPSS
  - Analysis
    - Scale
      - Reliability analysis
        » Statistics
        » Item
        » Scale if item deleted
Validity

- Face validity
- Content validity
- Predictive validity
- Construct validity
- Differential validity
- Consequential validity
Standardization

- Calculate means and standard deviation of norm group
- Provide norm table or conversion
  - Standard scores $z = \frac{x - \text{mean}}{\text{s.d.}}$
  - Standardised scores
    - T-scores = $z \times 10 + 50$
    - Stanine = $z \times 2 + 5$ (min = 1, max = 9)
    - Sten = $z \times 2 + 5.5$ (min = 1, max = 10)
    - IQ format = $Z \times 15 + 100$
Freedom from Bias

- Item bias
- Intrinsic test bias
- Extrinsic test bias
- Adverse impact
- Equivalence
- Differential Item Functioning
Writing the handbook

• Include copyright notice
• Include the scoring key and instructions
• Give evidence of reliability and validity
• Provide norms