



The Psychometrics Centre

RDI Summer School

2: Test construction

Professor John Rust
University of Cambridge



The Psychometrics Centre

Constructing a psychometric test

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



The Psychometrics Centre

Designing the blueprint

- Curriculum based
 - Bloom's taxonomy of educational objectives
- Job description
 - The job analysis
 - The person specification
- Theoretical
 - Ability
 - Personality

Knowledge Test Specification

		Arithmetic	Geometry	Algebra	Statistics
Manifestations	Knowledge of Terms (25%)	4	4	4	4
	Understanding (25%)	4	4	4	4
	Application (25%)	4	4	4	4
	Generalisation (25%)	4	4	4	4

Personality Test Specification

Content areas					
Manifestations		Extraversion	Neuroticism	Detail	Tough-mindedness
	High/ Positive	4	4	4	4
	High/ Negative	4	4	4	4
	Low/ Positive	4	4	4	4
	Low / Negative	4	4	4	4

UNIVERSITY OF
CAMBRIDGE
The Psychometrics Centre

What can be wrong with multiple choice items?

- Language
 - appropriateness
 - understandability
 - discontinuity
- Order
- Distractors
 - distractor analysis
- Bias

UNIVERSITY OF
CAMBRIDGE
The Psychometrics Centre

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

The item response table

	a	b	c	d	e	
1	4	2	3	1	3	
2	4	2	3	2	4	
3	4	2	3	2	1	
4	4	2	4	2	1	
5	4	4	2	1	3	
correct	4	2	3	1	2	

The item analysis table

	a	b	c	d	e	
1	1	1	1	1	0	
2	1	1	1	0	0	
3	1	1	1	0	0	
4	1	1	0	0	0	
5	1	0	0	1	0	

The item analysis table

	a	b	c	d	e	Score
1	1	1	1	1	0	4
2	1	1	1	0	0	3
3	1	1	1	0	0	3
4	1	1	0	0	0	2
5	1	0	0	1	0	2

Scoring items not people

	a	b	c	d	e	Score
1	1	1	1	1	0	4
2	1	1	1	0	0	3
3	1	1	1	0	0	3
4	1	1	0	0	0	2
5	1	0	0	1	0	2
Item score	5	4	3	2	0	

The difficulty value

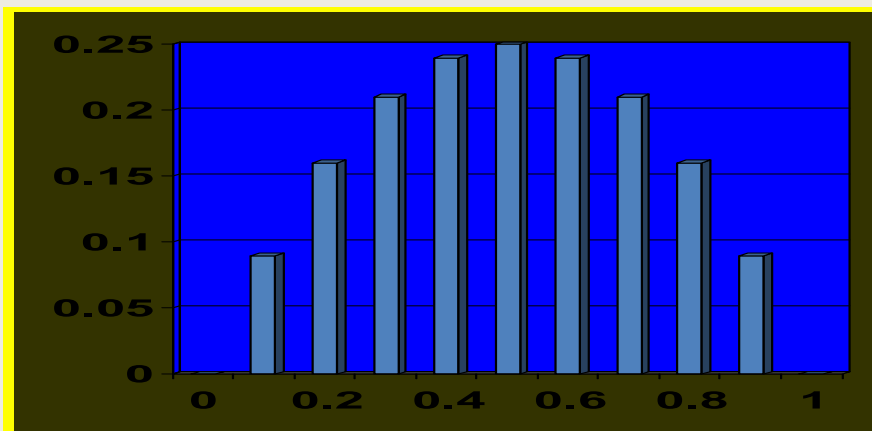
	a	b	c	d	e	Score
1	1	1	1	1	0	4
2	1	1	1	0	0	3
3	1	1	1	0	0	3
4	1	1	0	0	0	2
5	1	0	0	1	0	2
Item score	5	4	3	2	0	
p	1.0	0.8	0.6	0.4	0.0	

Contribution to total variance

	a	b	c	d	e	Score
1	1	1	1	1	0	4
2	1	1	1	0	0	3
3	1	1	1	0	0	3
4	1	1	0	0	0	2
5	1	0	0	1	0	2
Item score	5	4	3	2	0	
p	1.0	0.8	0.6	0.4	0.0	
p(1-p)	0.0	0.16	0.24	0.24	0.0	

UNIVERSITY OF
CAMBRIDGE
The Psychometrics Centre

P vs p(1-p)



UNIVERSITY OF
CAMBRIDGE
The Psychometrics Centre

Item discrimination

	a	b	c	d	e	Score
1	1	1	1	1	0	4
2	1	1	1	0	0	3
3	1	1	1	0	0	3
4	1	1	0	0	0	2
5	1	0	0	1	0	2
pbi	-	0.53	0.87	0.22	-	

Point biserial correlation

- $Pbi = ((x - \text{mean}) \sqrt{p(1-p)})/s.d.$
- Contribution of intercorrelations to the total variance
 - From the item variances
 - From the Item covariances

Using difficulty (p) and discrimination (r) indices

- p should be between .2 and .8
- r should be above approx .2
- Remember the test specification!

Software

- SPSS
 - Analysis
 - Scale
 - Reliability analysis
 - » Statistics
 - » Item
 - » Scale if item deleted

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%