Is aberrant response behaviour an inherent characteristic of students taking classroom maths tests?

Dr Panayiotis Panayides

In the context of Rasch measurement an aberrant response pattern in a test is one that is improbable, given that the data fit the Rasch model.

Types of individuals whose response patterns do not fit the typical pattern

SLEEPERS get bored and do poorly on the last items
(Linacre & Wright, 1994; Molenaar & Hoijtink, 1996)

FUMBLERS get confused with the item format and do poorly in the beginning of the test (Bracey & Rudner, 1992)

PLODDERS take too much time to answer and never get to the later items (Meijer 1996; Wright, 1977)

GUESSERS AND CHEATERS
(Athanasou and Lamprianou, 2002; Rudner, 1983).

People who show ***extreme creativity*** in interpreting questions (Karabatsos, 2000; Meijer, 1996) or with ***poor language skills***.

Possible factors associated with aberrance (as reported in the literature)

Gender
Mismatch between curriculum and test content
Position on the ability/trait scale
Test anxiety
Motivation
Class effect

<u>Panayides (2009)</u>
Attention Deficit Hyperactivity Disorder (ADHD)
Maths self-esteem
Language competency

Smith (1986) and Lamprianou (2005) suggested that an individual with an aberrant response pattern may exhibit such response behaviour in other testing situations too, implying that misfit could be an inherent characteristic of individuals.

The study

Test 1 (Diagnostic)

27 items (1 – 5 marks)
3 multiple - choice items

3 schools
13 teachers
25 classes
635 students

Test 2 (on quadratic equations)

16 items
12 multiple - choice items (1 mark)
4 multistep items (4 marks)

3 schools
9 teachers
18 classes
445 students

Rasch Analyses

Rasch Partial Credit Model

Infit and outfit mean square statistics

Cut – off values for the mean square statistics

Items: 1.3

Students: 1.3, 1.4, 1.5, 1.6, 1.8, 2.0

RESULTS

Test 1 calibrations

Two misfitting items (outfit > 1.5) and three slightly misfitting (1.3 < outfit < 1.5)

19 students with infit and/or outfit > 2.7 were removed and data calibrated again

2nd calibration: Three items only slightly misfitting
Investigation: Too important NOT REMOVED

Summary of the results of the Rasch analyses

	N	Estimate of mean (SD)	Range	Reliab.	Separ. Index	Infit msq mean (SD)	Outfit msq mean (SD)
Students	616	1.03 (1.17)	-2.63 to 3.64	0.87	2.61	1.06 (0.40)	0.97 (0.47)
Items	27	0.0 (1.21)	-2.18 to 1.75	0.99	11.40	1.01 (0.15)	0.97 (0.25)

STANDARDIZED RESIDUAL variance (in Eigenvalue units)

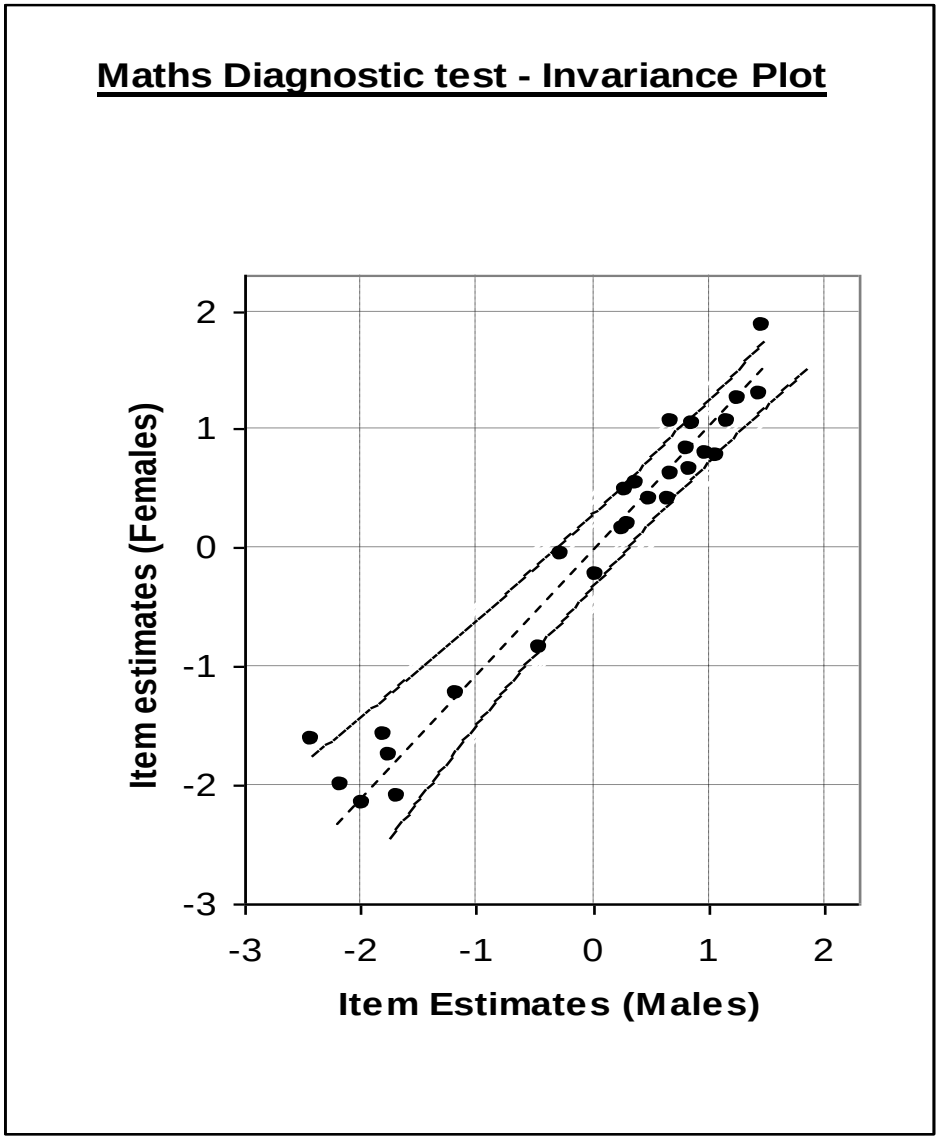
		Empirical			Modeled
Total variance in observations	=	61.7	100.0%		100.0%
Variance explained by measures	=	34.7	**56.3%**		59.1%
Unexplained variance (total)	=	27.0	43.7%	100%	40.9%
Unexpl var explained by 1st factor	=	2.0	**3.2%**	7.4%	

Variance explained by measures : variance explained by 1st factor = 17.4 : 1

Rules of thumb for the existence of a second dimension

1. **Eigenvalues < 2 : the implied dimension in the data has less than the strength of two items, and so, however powerful it may be diagnostically, it has little strength in the data.**
   Eigenvalue < 3 (in a reasonable length test) then the test is probably unidimensional. (Linacre, 2005)
2. **The first factor must explain a significant % of the unexplained variance (more than 20%)**
3. **A significant % of the total variance in the data**
   (Linacre, 2005, eigenvalue 2.7, N = 14, 0.2% of total variance)

Invariance plot for Test 1(Item calibrations from male and female groups)

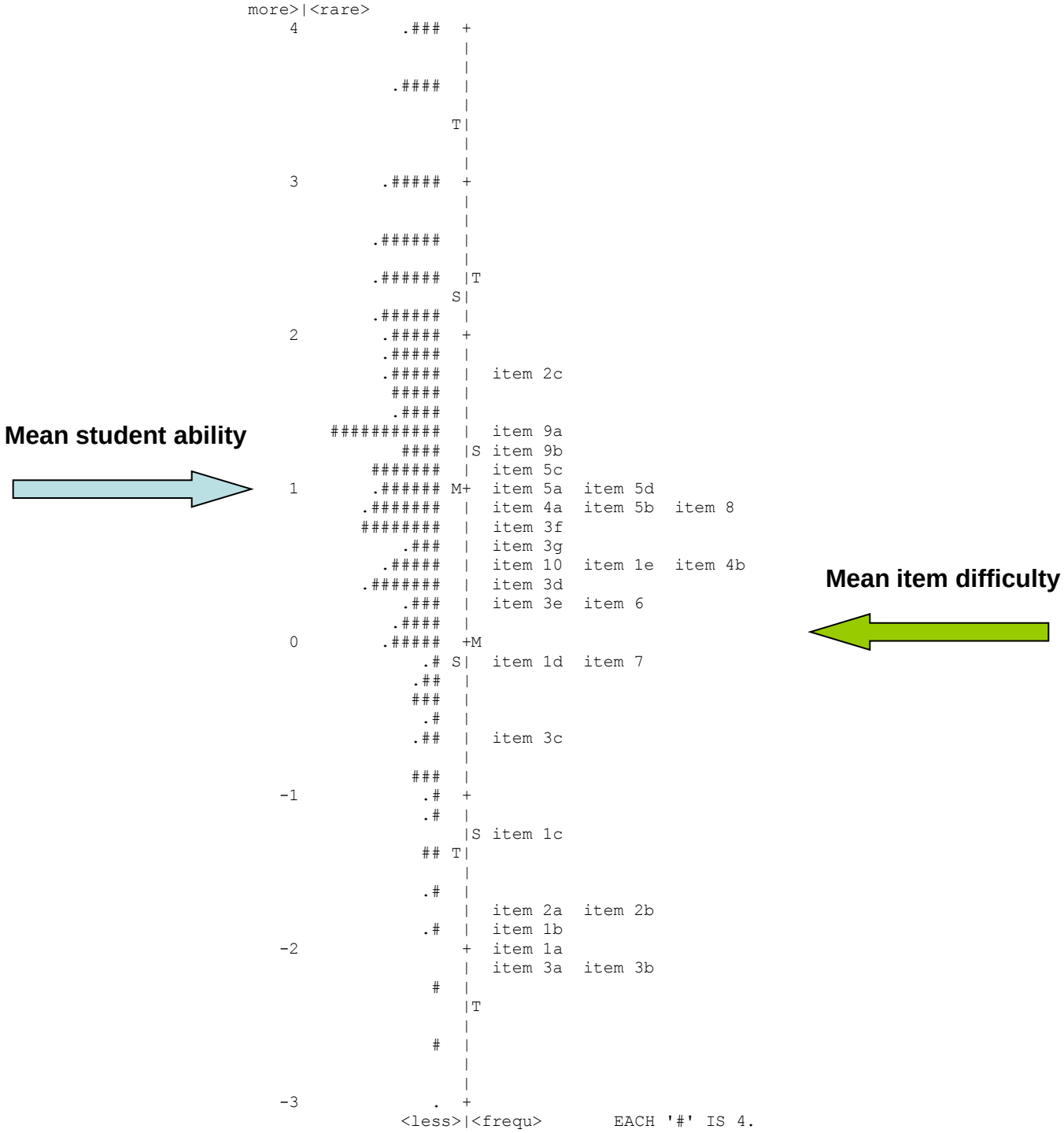


r = 0.975

Correlations of test scores with final exams

School 1: r = 0.795 (N = 287)
School 2: r = 0.704 (N = 37)
School 3: r = 0.701 (N = 281)

All ($p < 0.01$)


more>|<rare>
4 .### +
.####
T|
3 .##### +
.######
.######
S|
.######
2 .##### +
.#####
.#####
item 2c
#####
.####
########## item 9a
|S item 9b
####### item 5c
1 .###### M+ item 5a item 5d
.####### item 4a item 5b item 8
######## item 3f
.### item 3g
.##### item 10 item 1e item 4b
.####### item 3d
.### item 3e item 6
.####
0 .##### +M
.# S| item 1d item 7
.##
###
.#
.## item 3c
###
-1 .# +
.#
|S item 1c
T|
.#
item 2a item 2b
.# item 1b
-2 + item 1a
item 3a item 3b
#
|T
#
-3 . +
<less>|<frequ> EACH '#' IS 4.
Mean student ability
Mean item difficulty

Test 2 calibrations

Only one misfitting item (Item 13, difficulty – 1.07, outfit = 1.78)

6 high scorers (estimates 1.12 – 3.21) lost one or two marks

When removed  **outfit = 1.24**

Summary of the results of the Rasch analysis for Test 2

	N	Estimate of mean (SD)	Range	Reliab.	Separ. Index	Infit msq mean (SD)	Outfit msq mean (SD)
Students	445	0.25 (1.29)	-3.30 to 3.21	0.82	2.13	0.96 (0.67)	1.08 (0.79)
Items	16	0.0 (1.13)	-2.09 to 1.68	0.99	10.32	0.99 (0.08)	1.08 (0.23)

STANDARDIZED RESIDUAL variance (in Eigenvalue units)

		Empirical			Modeled
Total variance in observations	=	46.9	100.0%		100.0%
Variance explained by measures	=	**30.9**	**65.9%**		65.3%
Unexplained variance (total)	=	16.0	34.1%	100%	34.7%
Unexpl var explained by 1st factor	=	**1.5**	**3.3%**	9.5%	

Variance explained by measures : variance explained by 1st factor = 20.6 : 1

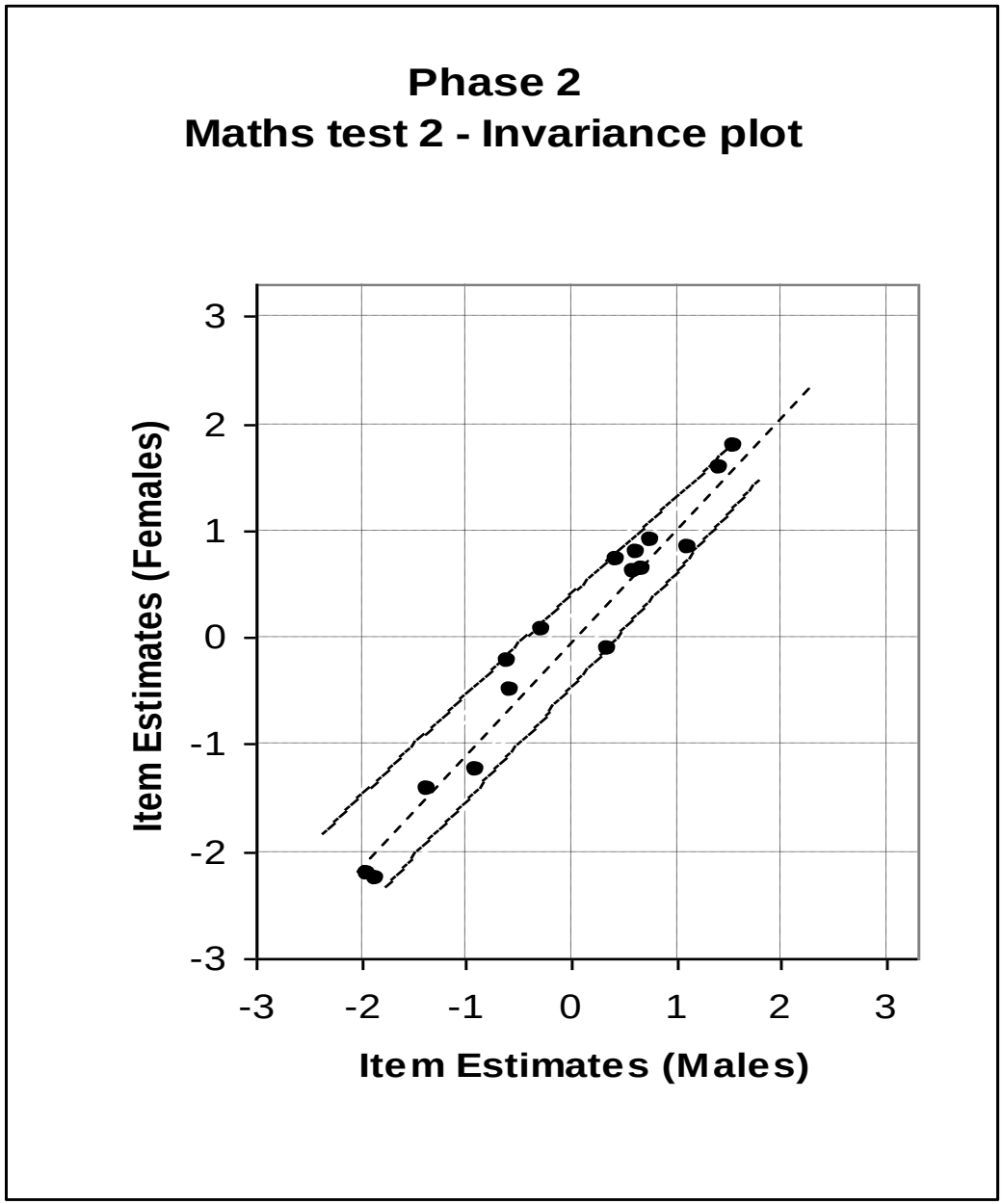
Correlations of test scores with final exams

School 1: r = 0.840 (N = 259),
School 2: r = 0.634 (N = 36), (p < 0.01)
School 3: r = 0.751 (N = 141)

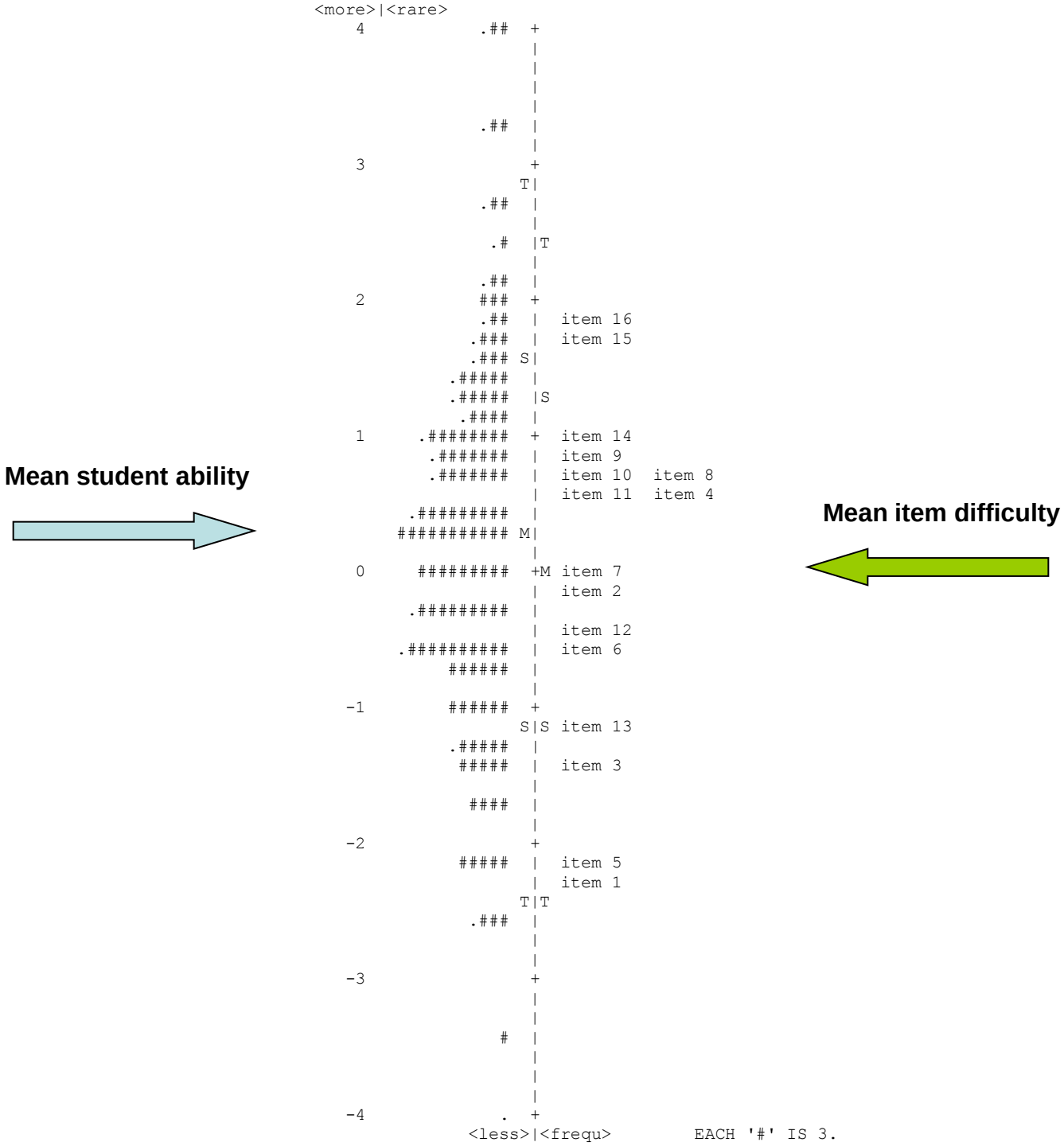
Table 4. Results of the analysis of the content validity questionnaire

Statements	Completely disagree	Disagree	Agree	Absolutely agree
The format of the questions is appropriate for the students	0	1	3	4
All the questions are clear and unambiguous	0	0	2	6
Students who know the answers have enough time to finish the test	0	2	4	2
All the important abilities and skills of the unit are assessed by the test	0	0	0	8
No irrelevant topics are included in the test	0	0	3	5
The test content is representative of the unit content as described in the curriculum	0	0	0	8

Invariance plot for Test 2 (Item calibrations from male and female groups)



r = 0.979


Mean student ability
Mean item difficulty
<more>|<rare>
4
3
2
1
0
-1
-2
-3
-4
item 16
item 15
item 14
item 9
item 10
item 8
item 11
item 4
item 7
item 2
item 12
item 6
item 13
item 3
item 5
item 1
<less>|<frequ>
EACH '#' IS 3.

Percentages of misfitting students for the various cut-off values.

		Test 1			Test 2		
		Outfit	Infit	Total	Outfit	Infit	Total
Cut-off values	1.3	22.8	23.3	35.0	23.6	21.6	30.8
	1.4	19.2	17.2	27.9	20.0	16.4	27.4
	1.5	15.6	13.9	23.9	17.3	14.2	23,6
	1.6	12.6	11.3	20.0	16.2	12.4	22.0
	1.8	8.5	6.0	12.8	13.3	8.8	17.3
	2.0	6.1	4.6	9.9	11.7	6.7	14.4

Cut-off	% of Misfitting students in Test 2 from: Fitting students in Test 1	Misfitting students in Test1	Chi-square	p-value
1.3	31.9	28.7	0.514 (0.371)	0.474 (0.542)
1.4	27.4	27.4	0.000 (0.000)	0.999 (1.000)
1.5	23.2	24.8	0.104 (0.036)	0.747 (0.849)
1.6	22.0	22.0	0.000 (0.000)	0.991 (1.000)
1.8	17.2	18.2	0.034 (0.000)	0.854 (1.000)
2.0	14.0	18.2	0.573 (0.281)	0.449 (0.596)

Concluding remarks

- This study reports that **misfit is not an inherent characteristic of students taking classroom maths tests.**

- This finding, together with Panayides' (2009) findings of no association between a large number of possible factors and misfit in the same setting, lead to the following intuitive conclusion: **In classroom maths tests, although misfits do occur, they do not predict misfits in other tests and are not dependent on psychological or demographic characteristics of the test-takers.**

- Therefore, high school maths teachers who test their students regularly should be aware that this kind of response behaviour does occur (perhaps leading to invalid estimates of their students' abilities) but should not be too concerned about it since they have many test results for their students and thus many ability estimates.

Investigating the Dimensionality of a primary maths test

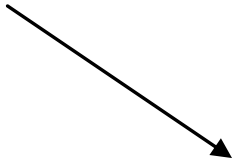
(Mis) Understanding the meaning of the equal sign

2. Circle the correct answer for the sum $27 + 46$

(α) 63 (β) 613 (γ) 73 (δ) 713

3. Match each operation in the first row with the appropriate answer in the second row, as shown in the example.

2 + 3 14 – 6 9 + 7 8 x 7



72 9 5 16 40 8 56

4. Complete the equations by filling in the blanks.

(a) 17 + 20 = …….

(b) ……. = 38 + 10

(c) 13 + 51 = 51 + …….

(d) 12 + ……. = 28 + 4

(e) 160 = ……. – 30

(f) 38 – 12 = ……. – 3

(g) …….. = 17

(h) 5 x 4 = ……. + 2

EFA

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	3,690	30,746	30,746	3,690	30,746	30,746	3,575
2	1,534	12,786	43,532	1,534	12,786	43,532	1,776
3	1,081	9,009	52,541	1,081	9,009	52,541	1,447
4	,985	8,210	60,751				
5	,824	6,867	67,619				
6	,764	6,369	73,988				
7	,680	5,665	79,653				
8	,621	5,177	84,830				
9	,583	4,858	89,688				
10	,491	4,091	93,779				
11	,446	3,717	97,496				
12	,301	2,504	100,000				

Extraction Method: Principal Component Analysis.

a. When components are correlated, sums of squared loadings cannot be added to obtain a total variance.

Principal Component Analysis

Items	Factor 1	Factor 2	Factor 3
2) 27 + 46 = …… (Multiple-choice)	,224	,062	**,663**
3a) 14 – 6 = ……. (Matching)	,136	**,681**	,300
3b) 9 + 7 = ……. (Matching)	,100	,208	**,810**
3c) 8 x 7 = ……. (Matching)	,074	**,694**	,165
4a) 17 + 20 = …….	,240	**,685**	-,139
4b) ……. = 38 + 10	**,503**	,375	,333
4c) 13 + 51 = 51 + …….	**,670**	,305	,137
4d) 12 + ... = 28 + 4	**,738**	,096	,153
4e) 160 = ……. – 30	**,705**	,147	,077
4f) 38 – 12 = ……. – 3	**,799**	,038	,140
4g) ……. = 17	**,622**	**,**218	,132
4h) 5 x 4 = ……. + 2	**,815**	,032	,132

Component Correlation Matrix

Component	1	2	3
1	1,000	,203	,178
2	,203	1,000	,160
3	,178	,160	1,000

Rasch Measurement

MOST MISFITTING RESPONSE STRINGS

student		**OUTMNSQ**	item	
			1	11
			352416197820	
		high		
18	1018	**9.90**	A|....0.....	**27 + 46 = …… (Multiple-choice)**
37	1037	**9.90**	B|...0.....	**8 x 7 = ……. (Matching)**
40	1040	**9.90**	C|...0.....	**8 x 7 = ……. (Matching)**
69	1069	**9.90**	D|.0.....	**17 + 20 =** …….
105	1105	**9.90**	E|....0.....	**27 + 46 = …… (Multiple-choice)**
119	1119	**9.90**	F|...0.....	**8 x 7 = ……. (Matching)**
121	1121	**9.90**	G|..0.....	**14 – 6 = ……. (Matching)**
178	1178	**9.90**	H|..00.....	**8 x 7 = ……. 14 – 6 = ……. (Matching)**
202	1202	**9.90**	I|..0.....	**14 – 6 = ……. (Matching)**

item STATISTICS: MISFIT ORDER

ENTRY NUMBER	RAW SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEA CORR.	EXACT MATCH OBS%	EXACT MATCH EXP%	item	G
6	127	162	-.02	.24	1.01	.1	**2.16**	2.4	A .54	87.7	85.1	**item 4b**	0
4	153	162	-2.32	.40	1.08	.4	**1.63**	1.0	B .31	95.7	95.1	**item 3c**	0
9	83	162	2.00	.20	1.06	.6	**1.50**	1.9	C .68	75.9	78.3	**item 4e**	0
1	151	162	-2.02	.37	1.18	.8	.92	.1	D .35	93.2	94.2	item 2	0
7	72	162	2.45	.20	1.09	.8	.99	.1	E .70	76.5	78.5	item 4c	0
3	160	162	-4.25	.76	1.03	.3	.12	-1.6	F .21	98.8	98.8	item 3b	0
8	68	162	2.61	.20	1.03	.3	.96	-.1	f .71	77.2	78.6	item 4d	0
11	121	162	.31	.23	1.00	.1	.85	-.4	e .60	84.6	83.5	item 4g	0
2	155	162	-2.68	.44	.88	-.3	.86	.1	d .33	96.9	96.1	item 3a	0
12	55	162	3.16	.21	.81	-1.9	.73	-.8	c .77	84.0	79.4	item 4h	0
10	39	162	3.88	.22	.78	-2.2	.60	-1.0	b .78	88.3	81.3	item 4f	0
5	157	162	-3.13	.51	.77	-.5	.26	-1.1	a .33	98.1	97.0	item 4a	0
MEAN	111.8	162.0	.00	.33	.98	-.1	.96	.0		88.1	87.2		
S.D.	43.4	.0	2.68	.17	.13	.9	.55	1.2		8.2	8.0		

4b) ……. = 38 + 10

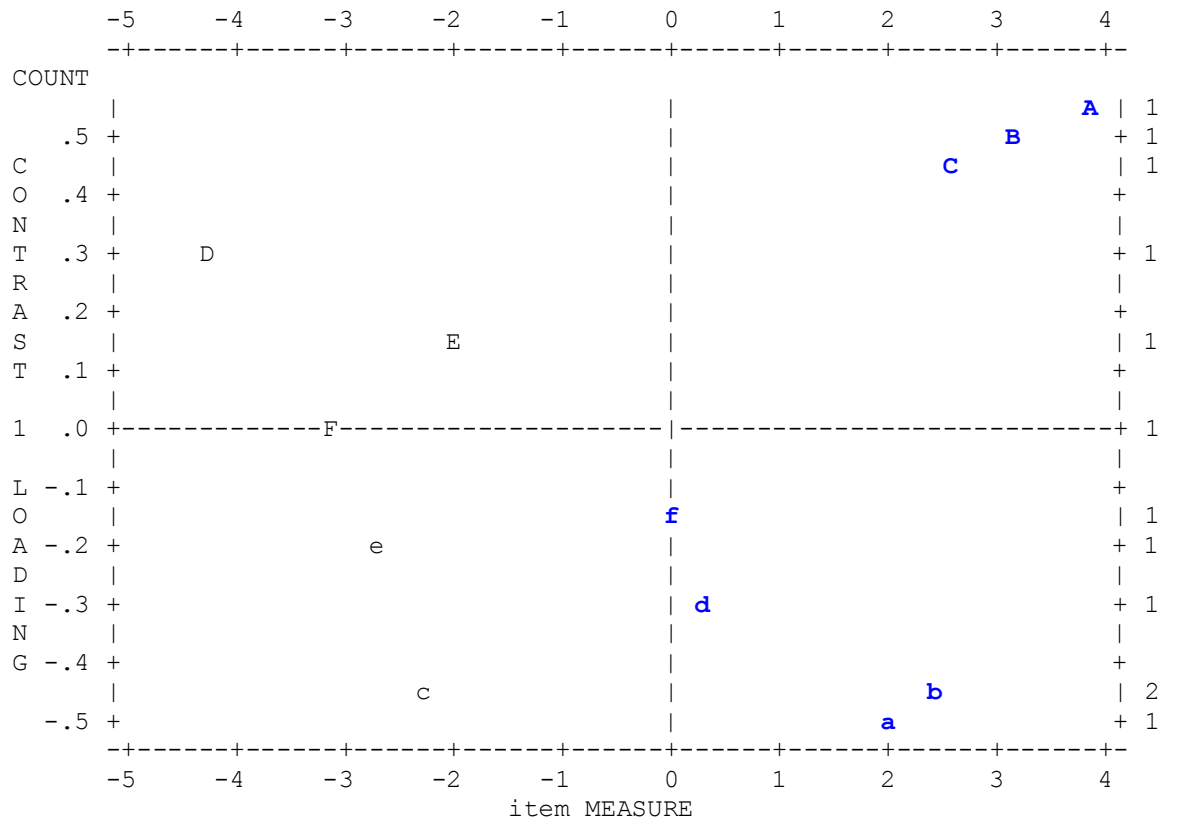
4e) 160 = ……. – 30

3c) 8 x 7 = ……. (Matching)

Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)

	-- Empirical --		Modeled
Total raw variance in observations =	29.6	100.0%	100.0%
Raw variance explained by measures =	**17.6**	**59.5%**	58.6%
Raw variance explained by persons =	7.8	26.2%	25.7%
Raw Variance explained by items =	9.9	33.3%	32.8%
Raw unexplained variance (total) =	12.0	40.5% 100.0%	41.4%
Unexplned variance in 1st contrast =	**1.6**	**5.6% 13.7%**	
Unexplned variance in 2nd contrast =	1.5	5.2% 12.8%	
Unexplned variance in 3rd contrast =	1.3	4.4% 10.9%	
Unexplned variance in 4th contrast =	1.2	3.9% 9.7%	
Unexplned variance in 5th contrast =	1.2	3.9% 9.6%	

STANDARDIZED RESIDUAL CONTRAST 1 PLOT



38 – 12 = ……. – 3

5 x 4 = ……. + 2

12 + ……. = 28 + 4

All positive estimates
The 7 most difficult

……. = 38 + 10

……. = 17

13 + 51 = 51 + …….

160 = ……. – 30

Remove the 9 highly misfitting pupils and do PCA again

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	3,855	32,122	32,122	3,855	32,122	32,122	3,630
2	1,524	12,698	44,820	1,524	12,698	44,820	2,102
3	1,076	8,965	53,784	1,076	8,965	53,784	1,549
4	,953	7,945	61,730				
5	,800	6,669	68,399				
6	,768	6,397	74,796				
7	,661	5,505	80,301				
8	,606	5,053	85,354				
9	,559	4,658	90,012				
10	,472	3,931	93,944				
11	,419	3,489	97,433				
12	,308	2,567	100,000				

Extraction Method: Principal Component Analysis.

a. When components are correlated, sums of squared loadings cannot be added to obtain a total variance.

	Factor 1	Factor 2	Factor 3
27 + 46 = …… (Multiple-choice)	,297	,130	**,630**
14 – 6 = ……. (Matching)	,219	**,682**	,368
9 + 7 = ……. (Matching)	,086	,218	**,843**
8 x 7 = ……. (Matching)	,162	**,701**	,228
17 + 20 = …….	,251	**,755**	-,128
……. = 38 + 10	**,475**	,463	,295
13 + 51 = 51 + …….	**,670**	,332	,198
12 + ……. = 28 + 4	**,741**	,165	,173
160 = ……. – 30	**,694**	,232	,088
38 – 12 = ……. – 3	**,809**	,097	,172
……. = 17	**,610**	,285	,126
5 x 4 = ……. + 2	**,816**	,118	,167

Component Correlation Matrix

Component	1	2	3
1	1,000	,277	,206
2	,277	1,000	,195
3	,206	,195	1,000