

Does Differential Item Familiarity Account for Differences in Performance Between TIMSS and PISA?

The Case of Mathematics and Science in England

Graham Ruddock,
National Foundation for
Educational Research

The Research Questions

- How familiar are pupils in England with the items used in TIMSS and PISA?
- Do differing levels of familiarity to pupils in England offer possible explanations for how well our pupils perform in TIMSS and PISA?



Methodology

- Rating Study with familiarity judged by experts.
- Based on earlier work on PISA 2000.

Methodology changed to allow for:

- 1) the inclusion of both PISA and TIMSS, 2 age groups, 2 different curricula
- 2) the overt differentiation in the curricula in England



Items rated for familiarity of:

- the concept or skill assessed
- the context
- item format
- overall appropriateness



The Rating Scale

	Not in Curriculum	Percentage of Students Estimated to be Familiar with the Concept/ Skill, Context/Text type, or Format/Task				
	None	Very few	Fewer than half	About half	More than half	All or almost all
	0%	about 5%	about 25%	about 50%	about 75%	about 95%
Code	0	1	2	3	4	5



Appropriateness Scale

	Do you feel that this question would be generally appropriate for your students?				
	Not at all appropriate				Very appropriate
Code	1	2	3	4	5



Items Reviewed

Mathematics: PISA mathematics literacy,
2002–2003

TIMSS mathematics grade 8,
1999 and 2003

Science: PISA science literacy,
2000 and 2003

TIMSS science grade 8,
1999 and 2003

English: PISA reading literacy,
2000 and 2003



Familiarity ratings for TIMSS and PISA science items

Science	Mean Score					Standard Deviation			Number of items	
	Is the Item Appropriate	Familiarity of				Is the Item Appropriate	Familiarity of			
		Concept	Context	Format		Concept	Context	Format		
TIMSS	3.36	3.07	3.07	2.81		1.38	1.55	1.46	1.45	268
PISA	3.23	3.06	3.05	2.60		0.86	0.84	0.79	0.84	45



Familiarity ratings for TIMSS and PISA mathematics items

Maths	Mean Score			Standard Deviation			Number of items		
	Is the Item Appropriate	Familiarity of Concept Context Format			Is the Item Appropriate	Familiarity of Concept Context Format			
TIMSS	4.42	3.55	4.17	3.65	0.55	0.74	0.42	0.80	278
PISA	3.49	3.46	3.69	2.92	0.87	0.90	0.72	0.83	97



Familiarity ratings for PISA reading literacy items

Reading	Mean Score			Standard Deviation			Number of items		
	Is the Item Appropriate	Skill	Text type	Format	Is the Item Appropriate	Skill		Text type	Format
PISA	3.54	3.65	3.51	3.47	0.55	0.51	0.55	0.57	144



PISA and TIMSS mathematics and science assessments compared

PISA	TIMSS		
<p style="text-align: center;"><i>Science Items</i></p> <p style="text-align: center;">PISA and TIMSS items seen as very similar in appropriateness and familiarity.</p> <p style="text-align: center;">In both, items rated as appropriate for over 50% of students, around 50% of students estimated to be familiar with the concept assessed and the item context. Item format estimated to be familiar to over 40% of students.</p>			
<p style="text-align: center;"><i>Mathematics Items</i></p> <p style="text-align: center;">TIMSS items seen as more appropriate and familiar in context and format than PISA items.</p> <p style="text-align: center;">The concept assessed in both PISA and TIMSS estimated to be familiar to around 60% of students.</p> <table border="0" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; border-right: 1px dashed black; padding-right: 10px;"> <p><i>PISA mathematics items</i></p> <p>Seen as appropriate for about 60% of students</p> <p>Context familiar to nearly 70% of students</p> <p>Item format familiar to nearly 50% of students</p> </td> <td style="width: 50%; padding-left: 10px;"> <p><i>TIMSS mathematics items</i></p> <p>Seen as appropriate for about 85% of students</p> <p>Context familiar to about 80% of students</p> <p>Item format familiar to about 65% of students</p> </td> </tr> </table>		<p><i>PISA mathematics items</i></p> <p>Seen as appropriate for about 60% of students</p> <p>Context familiar to nearly 70% of students</p> <p>Item format familiar to nearly 50% of students</p>	<p><i>TIMSS mathematics items</i></p> <p>Seen as appropriate for about 85% of students</p> <p>Context familiar to about 80% of students</p> <p>Item format familiar to about 65% of students</p>
<p><i>PISA mathematics items</i></p> <p>Seen as appropriate for about 60% of students</p> <p>Context familiar to nearly 70% of students</p> <p>Item format familiar to nearly 50% of students</p>	<p><i>TIMSS mathematics items</i></p> <p>Seen as appropriate for about 85% of students</p> <p>Context familiar to about 80% of students</p> <p>Item format familiar to about 65% of students</p>		



PISA and TIMSS mathematics and science assessments compared)

PISA	TIMSS
<i>Reading Demands</i> A lot of text to read, much more than in English practice; grammatical complexity also an issue.	<i>Reading Demands</i> Amount to be read usually similar to English practice, some awkward or complex phrasing.
<i>Subject Demands</i> The mathematics and science required often relatively low level as it is often being applied in a context new to the student.	<i>Subject Demands</i> The mathematics and science required often similar to practice in England.
<i>Use of Context</i> Usually large scale contexts with a number of marks related to them, but moderate number of contexts for students to deal with.	<i>Use of Context</i> Usually small scale contexts with one or two marks related to them, and, compared with key stage 3 tests, a large number of contexts for students to deal with.



Similarities and Differences between the Mathematics Assessments

	PISA	TIMSS	KS3	GCSE
In context	All questions	Some questions	Some questions	Some questions
Amount of Reading	Often Higher	Low	Low	Low
Complex sentence structures	Yes	Yes	No	No
2 page questions	Some	Some	Some	Some
Short questions	Some	Common	Some	Common
Multiple choice	Some	Over 50%	A few	A few



Similarities and Differences between the Science Assessments

- KS2 to KS3 to GCSE: mainly questions presented on one or two complete pages. In either case a scientific theme or context sustains a number of marks.
- Students taking PISA will also have seen shorter science questions in preparation for GCSE. Situation less clear for TIMSS.
- PISA science questions (2000 and 2003) are likely to be unfamiliar to English students because of the amount of reading required and their rather different structure.



Science: PISA or TIMSS for the best Performance?

- Equally familiar and appropriate in the rating exercise
- But they are very different.
- PISA Science large scale contexts, more reading than in England.
- TIMSS short items, whereas 1 or 2 page questions predominate in England.
- More contexts to deal with in TIMSS.
- How do these affect performance?



Mathematics: PISA or TIMSS for the best Performance?

- TIMSS rated as more familiar and more appropriate than PISA.
- Both PISA and TIMSS use a range of item types, eg short response to multi-page questions. The balance differs.
- All of PISA in context
- How do these affect performance?

