



***Effects of student reading proficiency on
PISA mathematics items – A DIF study
across Countries***

Luc T. Le & Ray J. Adams

Structure of the presentation

- Background of DIF analysis
- Study design
- Results/findings

Differential Item Functioning

- DIF: examinees of different groups have different probabilities of answering items correctly *after matching on the underlying ability*.

Types of DIF

- *Uniform DIF*: The probability of answering an item correctly is greater for one group uniformly over all ability levels
- *Non-uniform DIF*: Probability of answering the item correctly is not greater across all levels of ability for any group

Detecting Method

- *Observed score base (Traditional):*
 - p -value, ANOVA - **Angoff (1972)**
 - MH - **Mantel & Haenszel (1959),
Holland & Thayer (1988)**
 - Log Regression - **Swaminthan & Roger (1990)**
- *Factor analysis base:* **Oort (1992), Muthén (1989)**
- *Latent base (IRT):* **Lord (1980), Raju (1998)**

Research questions

- *What are the relative weaknesses and strengths of the lower reading ability students on PISA mathematics items by PISA framework: format, content, context and competency?*
- *Is there any relationship between the uniform reading-related DIF and the item difficulty or between it and the item discrimination?*

Data

- 2003 PISA data
- Mathematics: student responses
- Reading: student WLE scores

Analysis design

- Grouping students in individual countries reading scores
- Calibrating items by the separated groups using PCM (Masters, 1982)
- Computing related DIF and flagging substantial DIF
- Exploring weakness and strength of Lower reading students by:
 - PISA frame work: item context, item content, competency, item format
 - Other factors: item difficulty, item discrimination, solving tasks in the items

Related DIF

- Related DIF measure: $DIF = d_F - d_R$

Where

d_F : item difficulty estimate from focal group (lower reading)

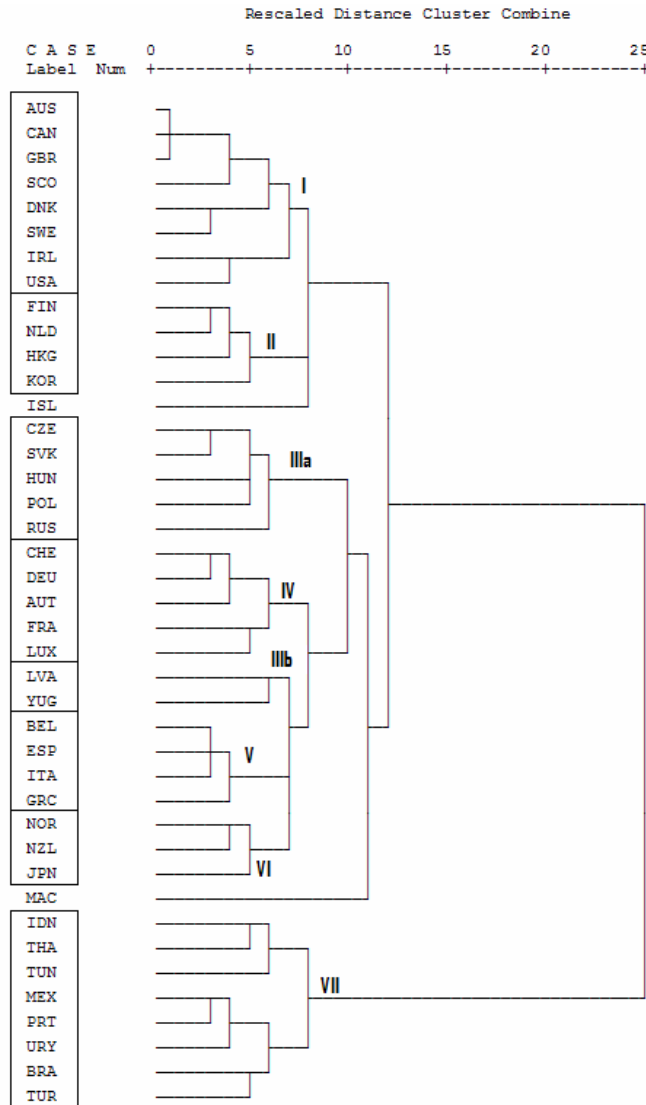
d_R : item difficulty estimate from reference group (higher reading)

- DIF >0: item is relatively harder for the lower reading
- DIF <0: item is relatively easier for the lower reading
- Flag defined: item substantially harder for lower reading students

DIF >0.25 and significant test of DIF <0.01

$$t = \frac{DIF}{\sqrt{SE_{d_F}^2 + SE_{d_R}^2}}$$

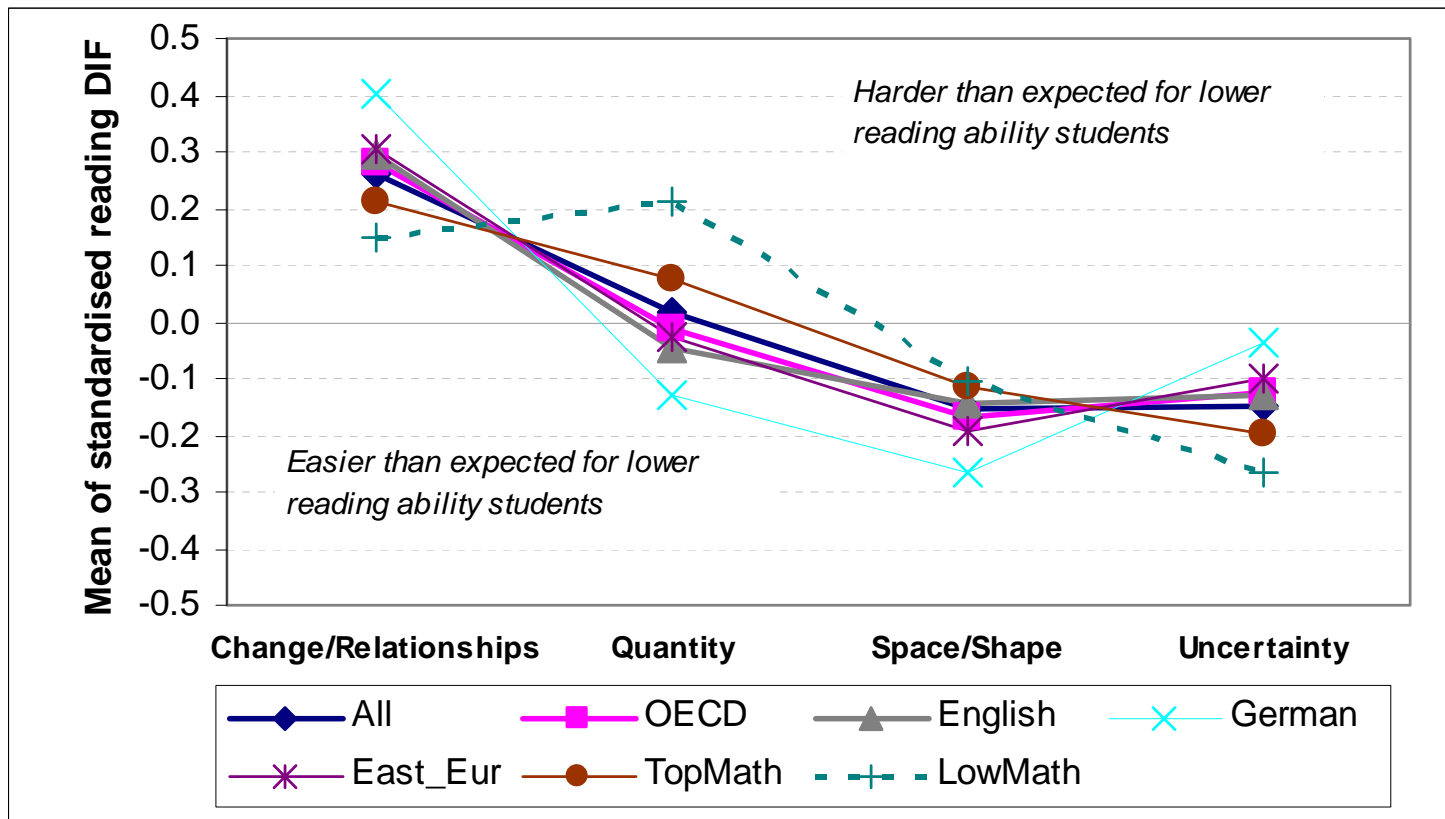
Cluster of mathematics reading-related DIF across countries



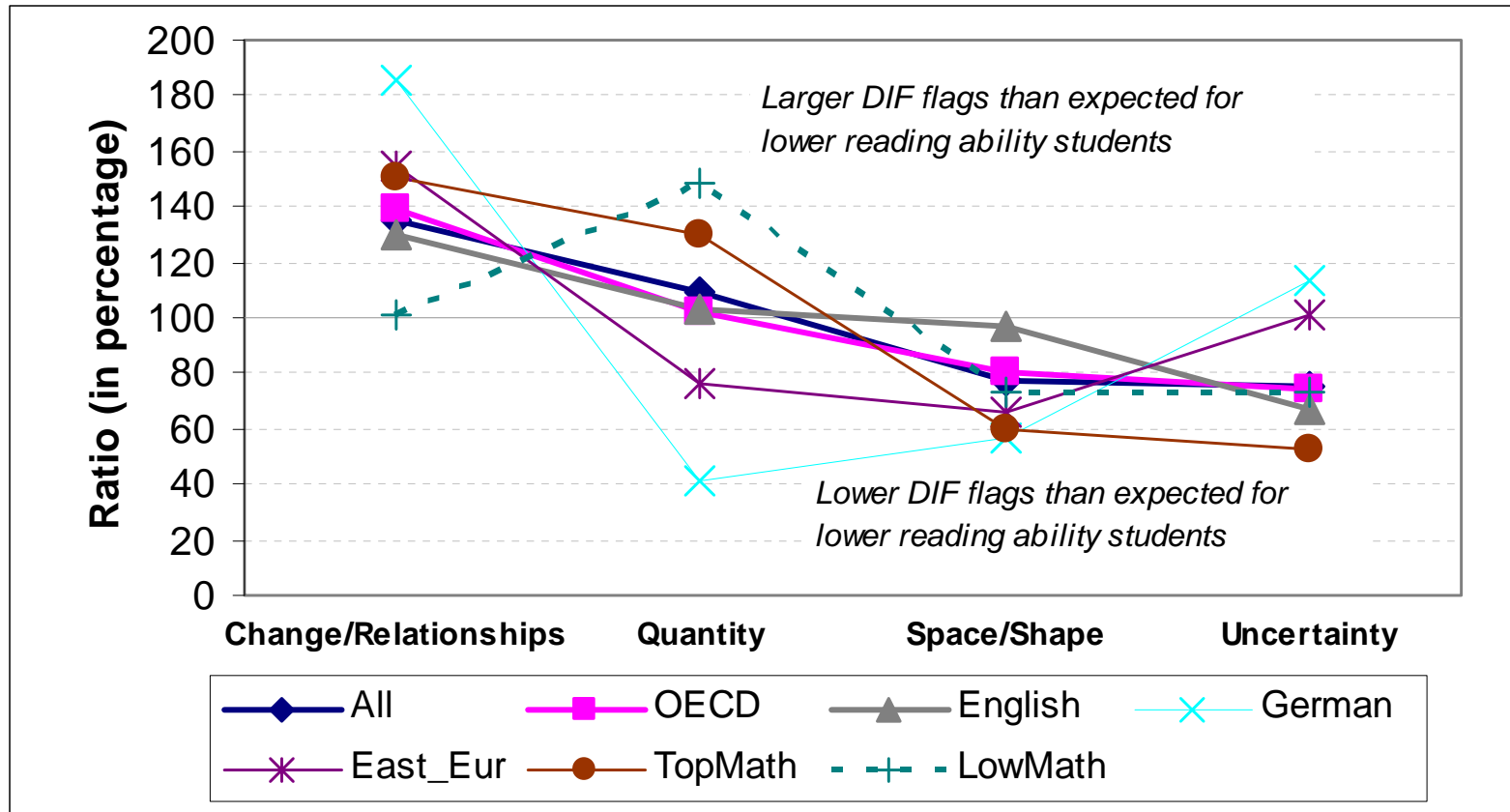
Five meaningful country clusters

- *English-speaking group*: Australia, New Zealand, Canada, England, Ireland, Scotland and United States;
- *German-speaking group*: Switzerland, Germany, Austria and Luxembourg;
- *Eastern European group*: Czech Republic, Slovak Republic, Hungary, Poland, Russia, Latvia and Serbia/Montenegro;
- *Top mathematics achievement group*: Finland, Netherlands, Hong Kong, Korea and Japan;
- *Low mathematics achievement group*: Indonesia, Thailand, Tunisia, Mexico, Uruguay, Brazil and Turkey.

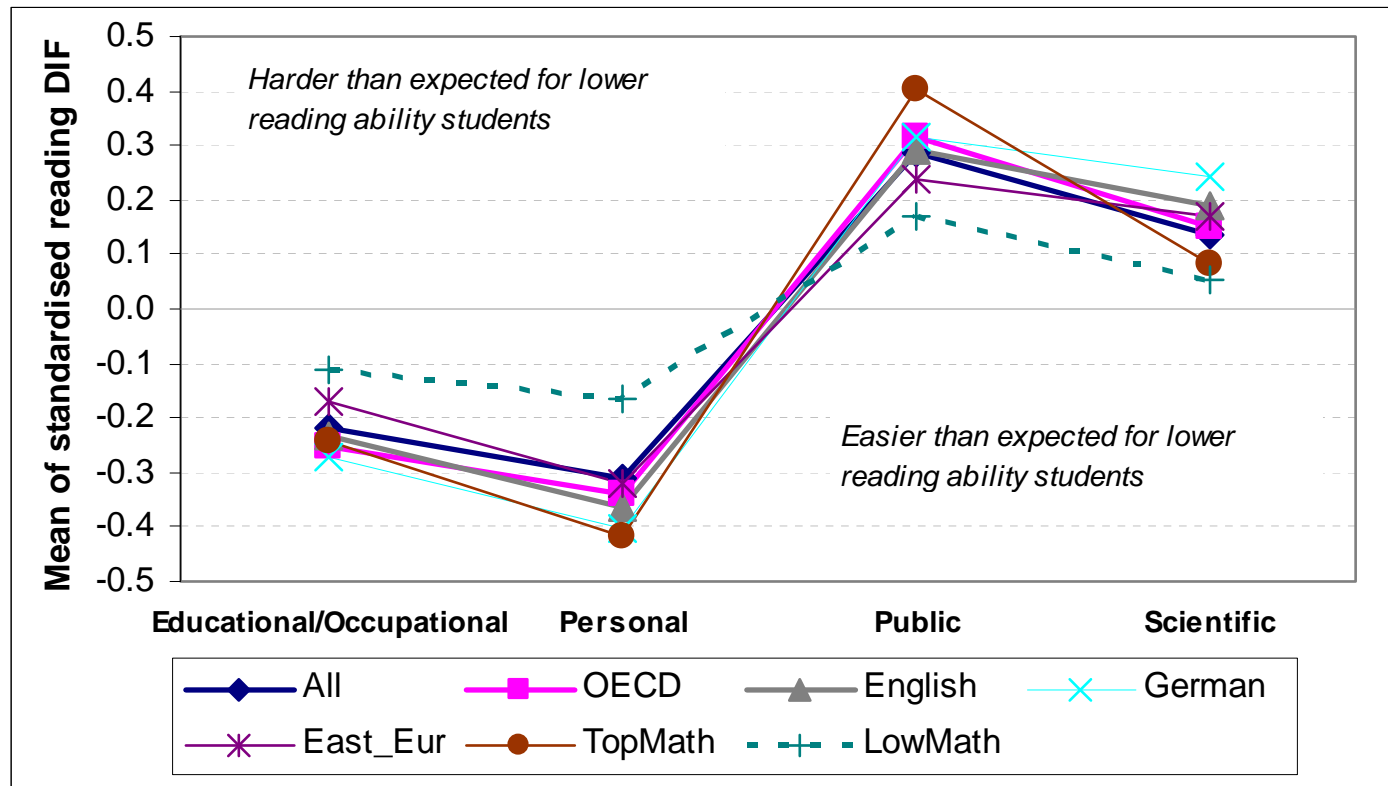
Mean of standardised reading-related DIF and item Content



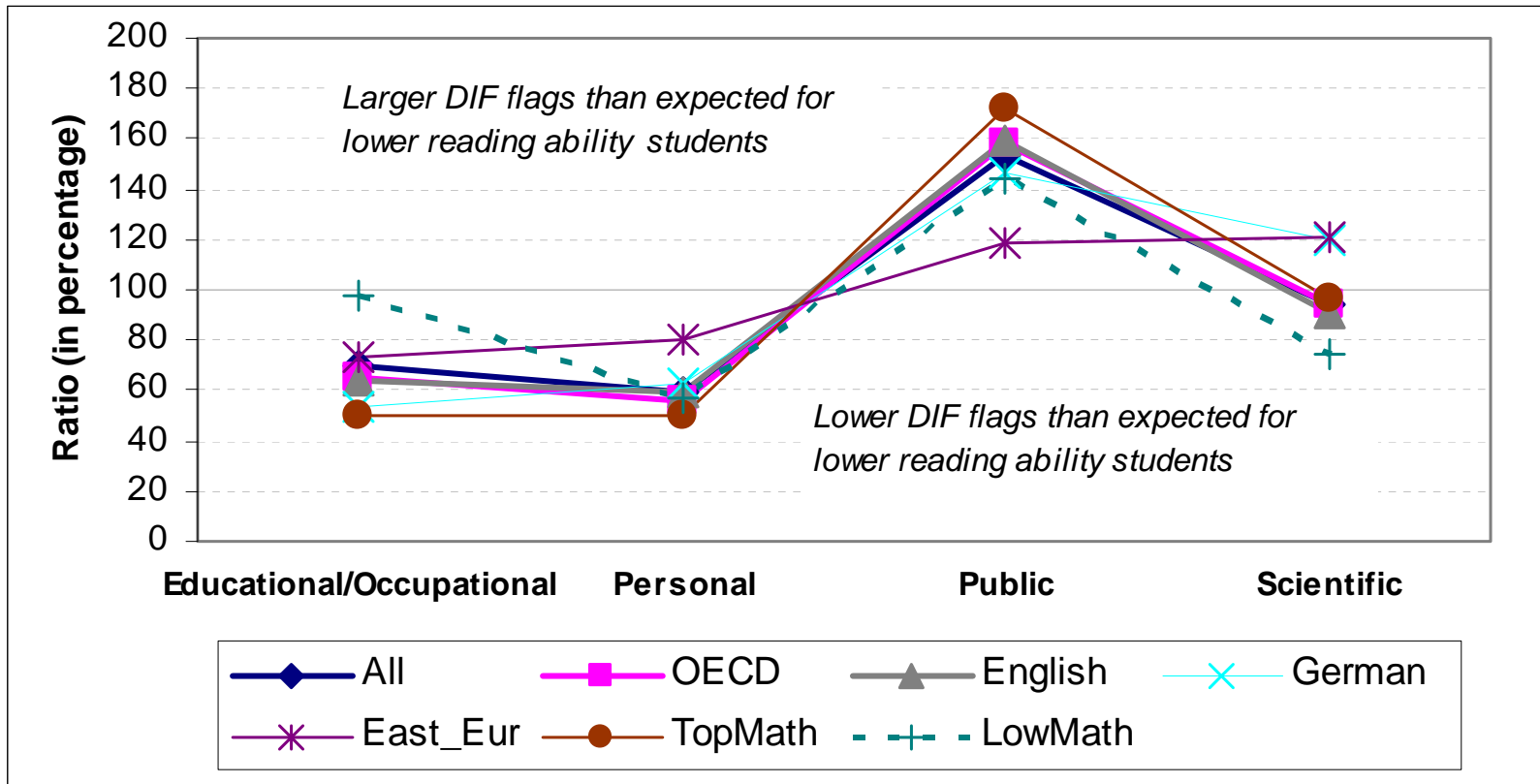
Ratio of reading-related DIF flags and item Content



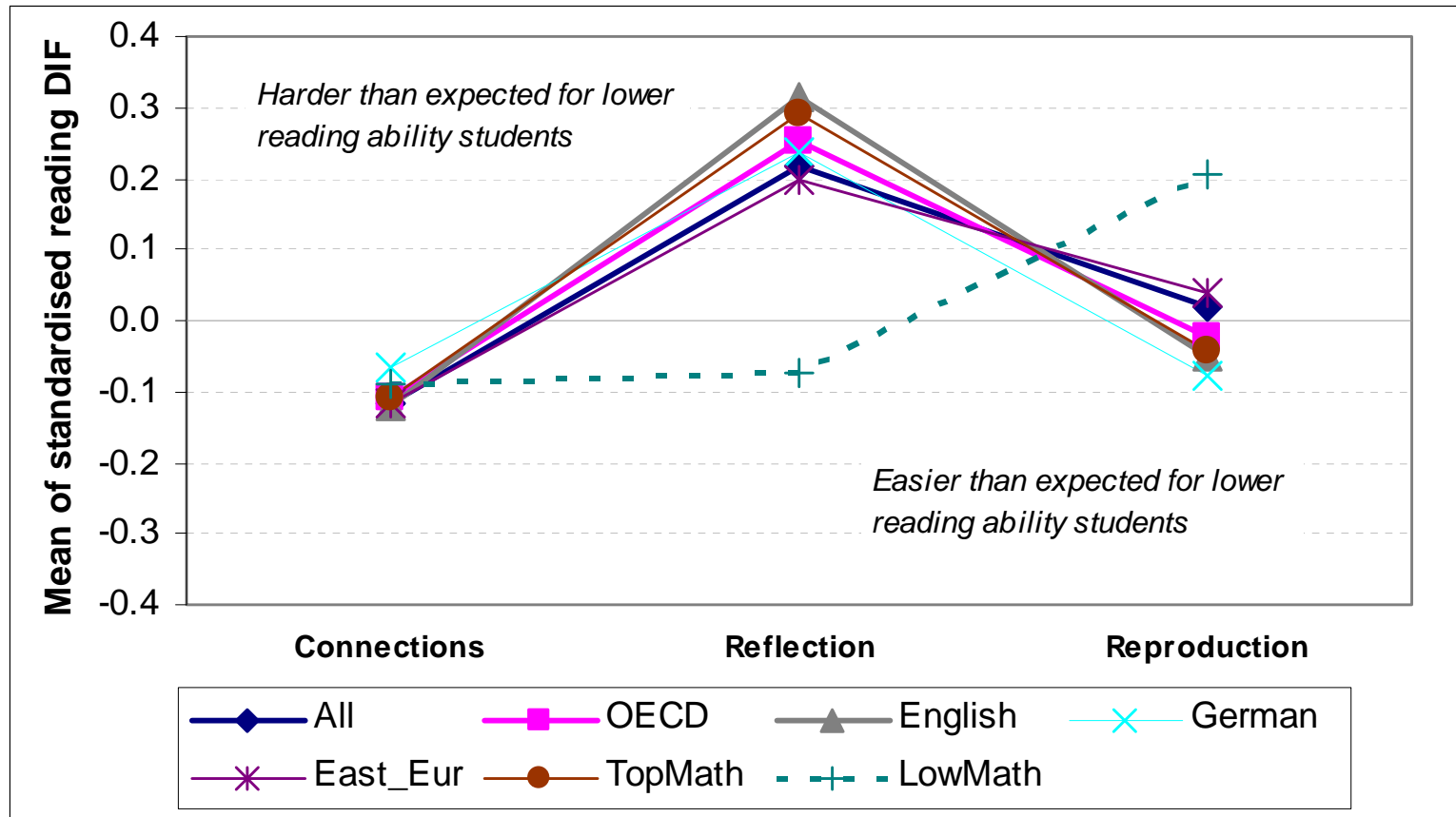
Mean of standardised reading-related DIF and item Context



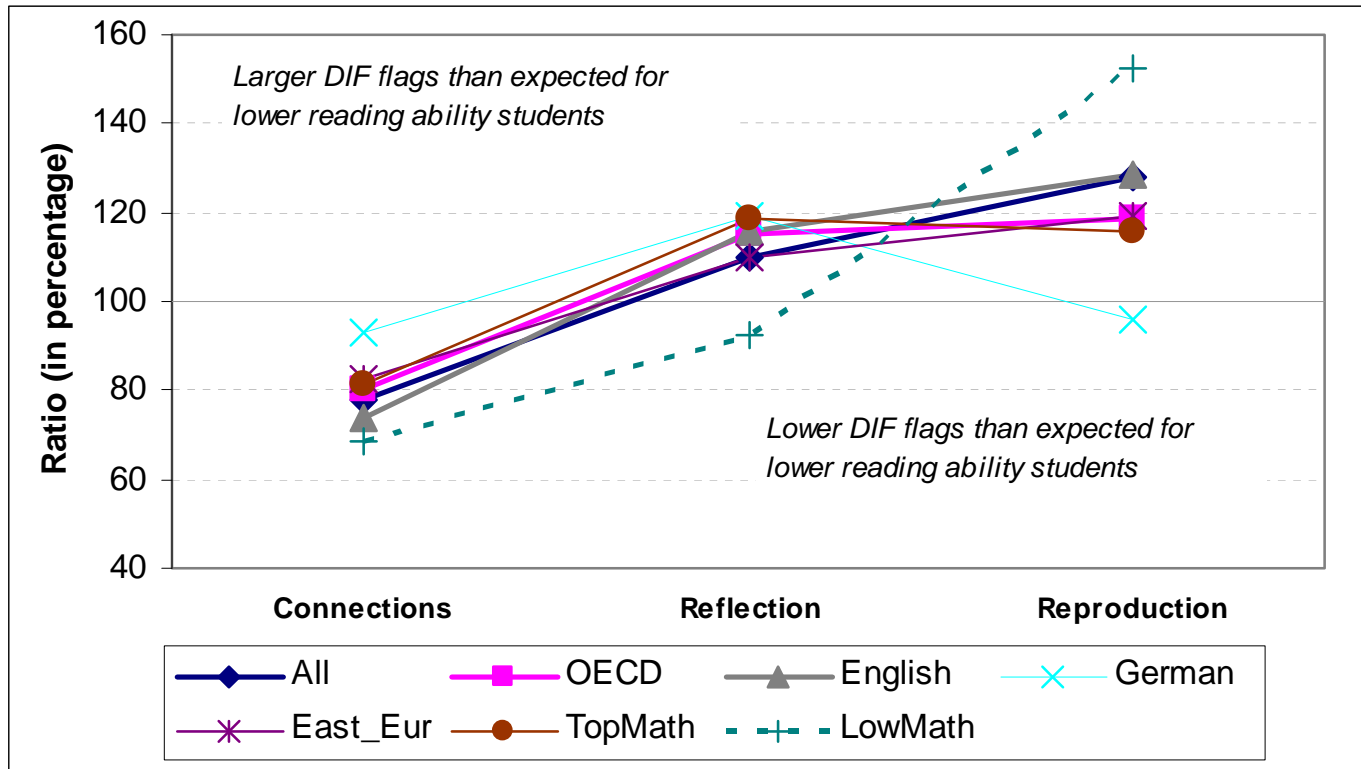
Ratio of reading-related DIF flags and item Context



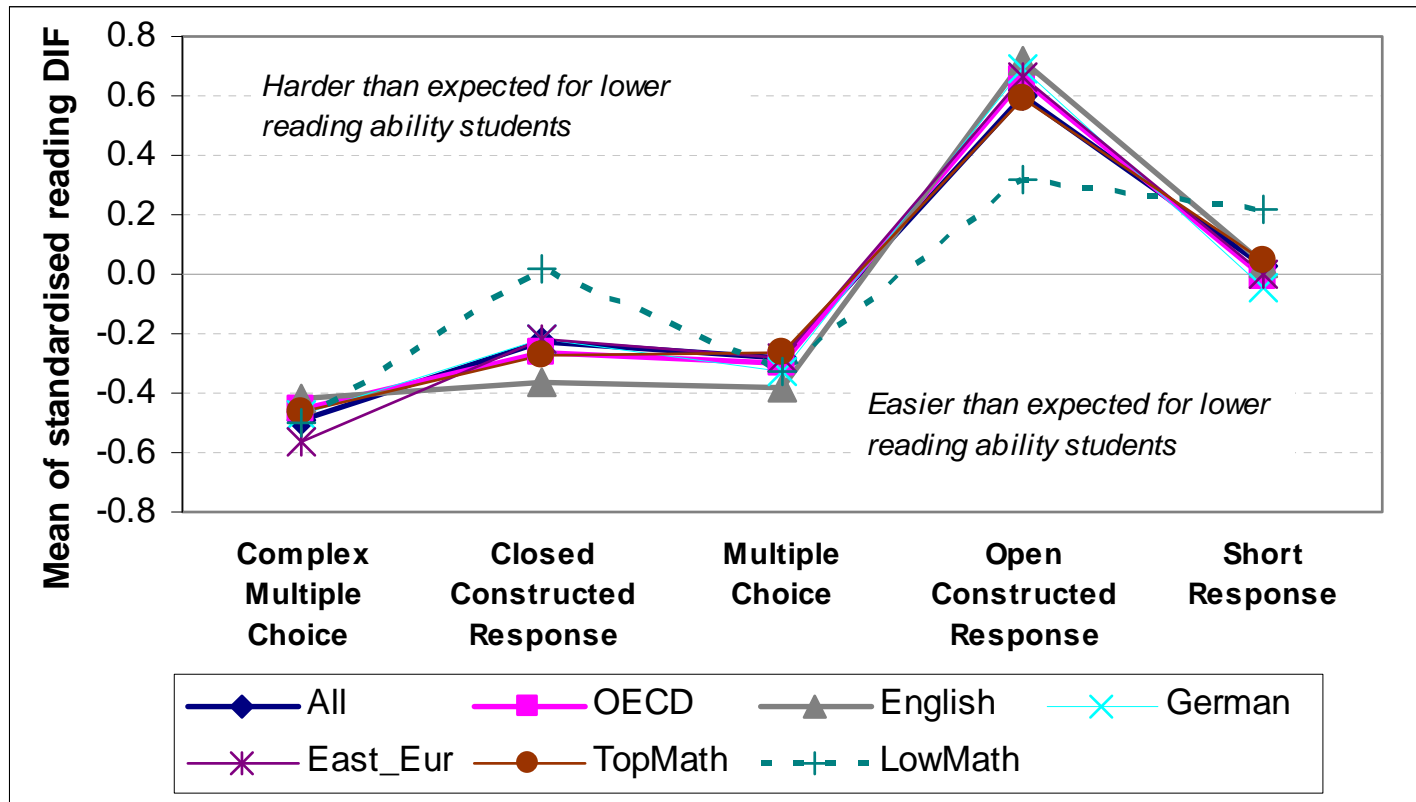
Mean of standardised reading-related DIF and item Competency



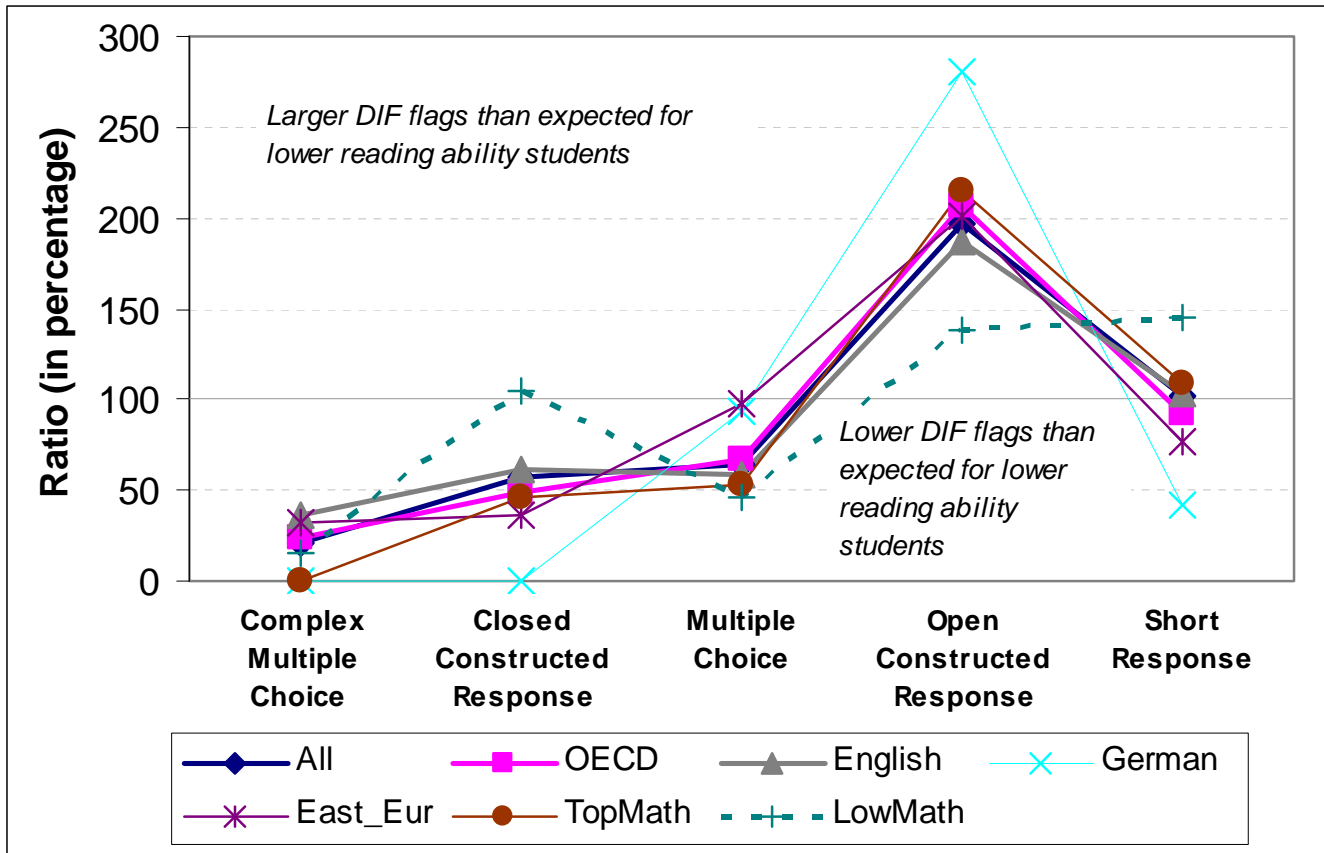
Ratio of reading-related DIF flags and item Competency



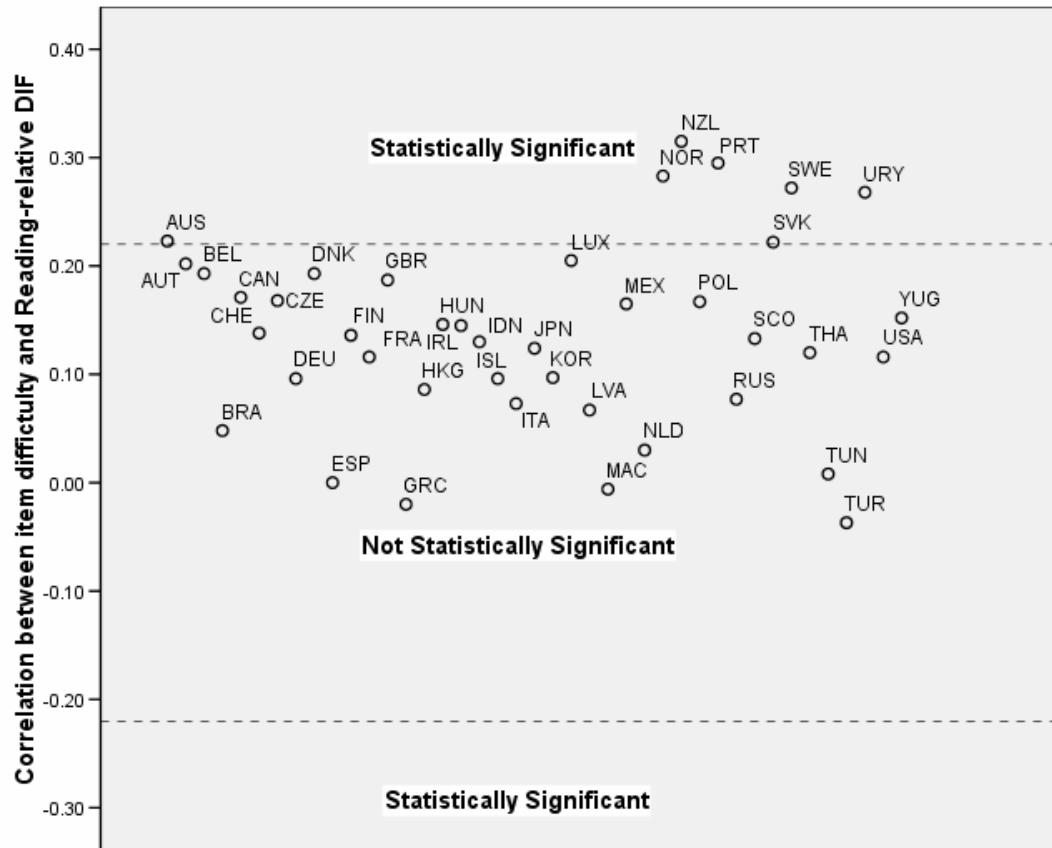
Mean of standardised reading-related DIF and item format



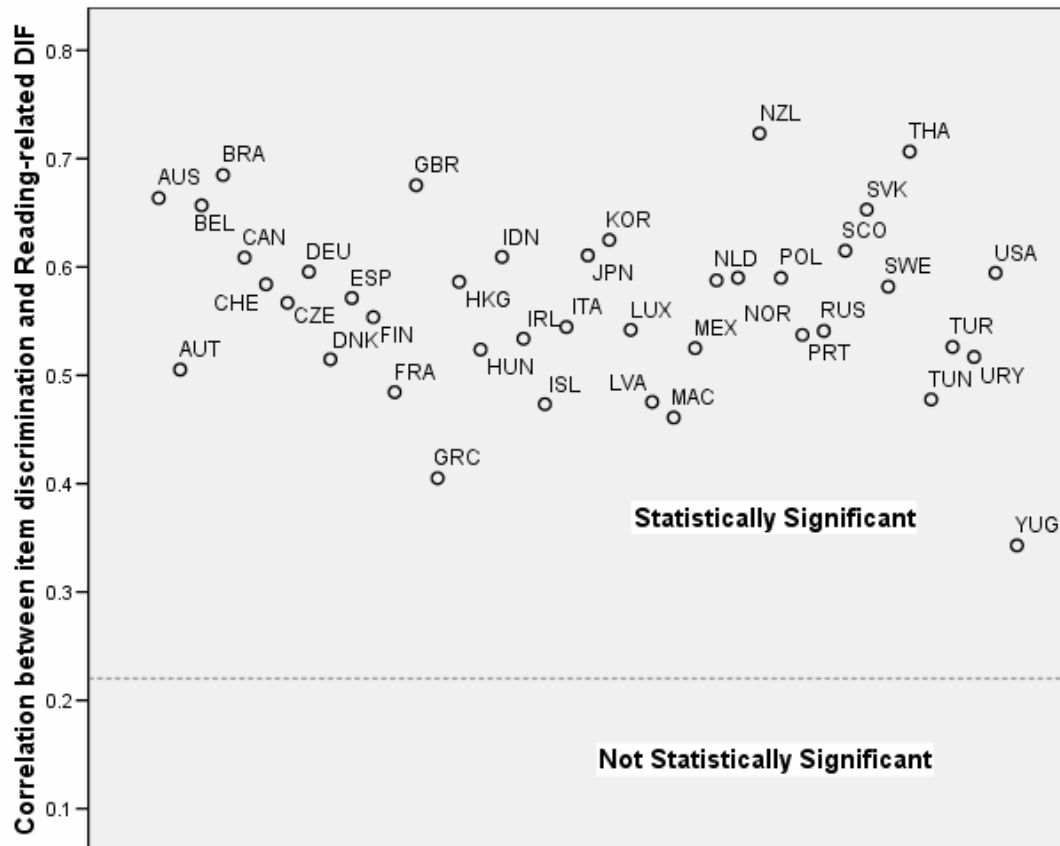
Ratio of reading-related DIF flags and item format



Correlation between reading-related DIF and item difficulty

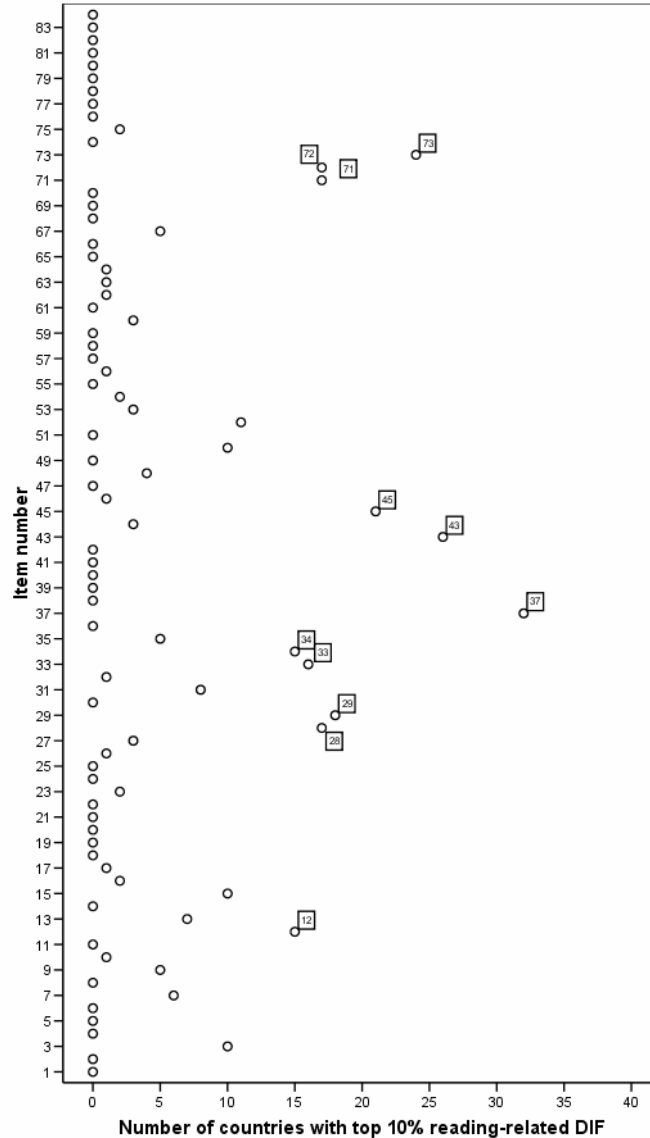


Correlation between reading-related DIF and item discrimination





Items with top 10% reading-related DIF by individual countries



Highest reading-related DIF items across countries

<i>Item No</i>	<i>Content</i>	<i>Context</i>	<i>Competency</i>	<i>Format</i>	<i>Country count</i>	<i>Per Cent</i>	<i>OECD Country count</i>	<i>Per cent</i>
37	Uncertainty	Educational/ Occupational	Reproduction	Open Constructed Response	32	78.0	25	80.6
43	Quantity	Public	Reflection	Closed Constructed Response	26	65.0	21	67.7
73	Change/ Relationships	Public	Reflection	Open Constructed Response	24	58.5	20	64.5
45	Change/ Relationships	Scientific	Reflection	Open Constructed Response	21	52.5	15	48.4
29	Space/Shape	Public	Reflection	Open Constructed Response	18	43.9	17	54.8
28	Space/Shape	Public	Connection	Open Constructed Response	17	41.5	15	48.4
71	Uncertainty	Public	Connection	Open Constructed Response	17	41.5	15	48.4
72	Change/ Relationships	Public	Reproduction	Short Response	17	41.5	14	45.2
33	Quantity	Public	Reproduction	Short Response	16	39.0	10	32.3
12	Change/ Relationships	Scientific	Connection	Open Constructed Response	15	36.6	13	41.9
34	Quantity	Public	Reproduction	Short Response	15	36.6	10	32.3

Lowest reading-related DIF Items

<i>Item No</i>	<i>Content</i>	<i>Context</i>	<i>Competency</i>	<i>Format</i>	<i>Country count</i>	<i>Per Cent</i>	<i>OECD Country count</i>	<i>Per cent</i>
20	Space/Shape	Educational/ Occupational	Connections	Complex Multiple Choice	31	75.6	24	77.4
40	Uncertainty	Personal	Reproduction	Multiple Choice	27	65.9	21	67.7
51	Quantity	Educational/ Occupational	Reproduction	Closed Constructed Response	18	43.9	16	51.6
24	Space/Shape	Public	Connection	Multiple Choice	16	39.0	15	48.4
25	Change/ Relationships	Personal	Connections	Short Response	15	36.6	13	41.9

Summary

- Different directions of relative weaknesses and strengths of the lower reading ability students in 4 dimensions: item context, content, competency, and format

Items harder than expected for lower reading students

- Items are in Open Constructed Response format
- Items contain a formula: understanding, applying or establishing a formula
- Items require selecting or ordering important information among number of different sources.



Items advantaged by lower reading students

- Items are in Multiple Choice format
- Items require a simple answer without explanation
- Items are based on few events or information
- Items require a simple computation or comparison
- Items contain a simple formula and require only a direct substitution

Relationship between related DIF and item difficulty

- The correlations of mathematics item difficulty and reading-related DIF are positive, but not significant
- Lower reading students perform differentially better on the easier items compared to matched higher reading students and differentially less well on the harder items.

Relationship between related DIF and item discrimination

- There is a significant correlation between reading-related DIF and item discrimination across the PISA countries
- Lower reading students perform differentially better on the lower discrimination items compared to matched higher reading students and differentially less well on the larger discrimination items.
- The DIF may be partly caused by the inconsistency between the item modelling (same slope) and the variation of their ICCs in slopes



Thank you