Gender differences in Variability and Extreme Scores

ECER, Berlin, 15 sept. 20121

Ariane Baye
Dominique Lafontaine
Christian Monseur

University of Liège Belgium

Gender differences in an international context

- International comparative surveys (such as IEA and PISA) of student achievement are useful tools for estimating the magnitude of the gender gap

Gender differences in an international context

- Review of literature
  - Reading: gender gap in favour of girls, at all ages, larger among the youngest students, both in achievement and attitudes
  - Physical sciences and Math: gender gap in favour of boys, both in achievement and attitudes, increasing with the age

Gender differences in an international context

- The magnitude of the gender gap may depend on:
  1. The test content
     1. Test content : reading, reflect upon and evaluate, spatial representation, physical science
     2. Test/Item format: smaller on non-continuous texts, females better on constructed-response
  2. The study design
     1. Target population: in math, differences only in high school
     2. Type of sample: representative sample vs. particular sample
  3. The statitics performed: effects sizes / central tendency statistics

Purpose of the study

- Exploring gender differences in reading, mathematics and science since the 1990s using international large-scale survey data (IEA and PISA)
- Hypothesis:
- (1a) gender differences at the extreme tails of the distribution may vary substantially compared to gender mean differences
- (1b) those differences will be in favour of males at the higher end of the distribution in mathematics and sciences, and to the disadvantage of males at the lower end of the reading distribution
- (2a) the supposed greater variability of males will vary according to the domain
- (2b) methodological choices such as the study design will influence the results

Method

- Data:
  - 6 IEA studies:
    - TIMSS 1995, 1999, 2003, 2007
    - PIRLS 2001, 2006
  - 4 PISA cycles:
    - PISA 2000, 2003, 2006, 2009 (every domain every cycle)

  => 1,393 cases, one case corresponding to an assessment of one domain in one population in one country

Method

- VARIABILITY
  - Gender Differences in Standard Deviation
    - Females SD – Males SD (negative and significant > males' sd is significantly higher than females' sd)
  - Gender Variance Ratio
    - Males variance/Females variance
- EXTREME SCORES
  - Gender differences at p. 5, 10, mean, 90, 95 with effect sizes
    - (Males mean- Female mean)/ Pooled standard deviation

Gender differences in variability

	N	Mal S	Mal NS	Fem NS	Fem S	VR
Reading	285	61%	34%	6%	0%	1.15
Math.	554	44%	47%	9%	0%	1.12
Sciences	554	48%	43%	8%	0%	1.14
Primary	327	33%	57%	9%	0%	1.11
Secondary	1066	54%	38%	8%	0%	1.14
IEA	779	37%	53%	11%	0%	1.12
OECD	614	63%	32%	4%	0%	1.16
Mean	1393	49%	43%	8%	0%	1.13

Results
Variability

- Gender Differences in Variability
  - Standards deviations: in 92% of the 1,393 cases, male standard deviation are larger than female ones
  - The difference is statistically significnt in 49% of cases
  - Females standard deviation is higher than males one in 8%, but this difference is significant for only 2 of the 1,393 cases
  - Greater SD for males varies according to the domain, education level, IEA and PISA

Results
Variability

- Gender Differences in Variability
  - Male/female variance ratio is 1.13 on average: male variance is 13% higher than female variance.

Results
Variability

- Variance:
  - Almost no exception to the higher male variance ratio
  - The differences between domains, educational levels, organisations and surveys are quite slight, except for the difference between PISA and PIRLS in reading

Results

Extreme scores

- Note: effect sizes, Cohen's (1977) categorisation:
  - 0.20-0.50: small
  - 0.50-0.80: medium
  - More than 0.80 : large

Gender diff in extreme scores

	N	P5	P10	Mean	P90	P95
Reading	285	**0.45**	**0.44**	**0.34**	**0.24**	**0.23**
Math.	554	0.03	0.02	-0.06	-0.13	-0.15
Sciences	554	0.03	0.02	-0.07	-0.15	-0.16
Primary	327	0.11	0.09	0.01	-0.05	-0.06
Secondary	1066	0.12	0.10	0.01	-0.07	-0.09
IEA	779	0.05	0.04	-0.04	-0.11	-0.12
OECD	614	**0.21**	**0.20**	0.10	0.00	-0.02
Mean	393	0.12	0.11	0.03	-0.05	-0.07

Positif Effect Size in favour of Females

Conclusion

- Extreme tails
- The gender differences at the extreme tails of the distribution are often more substantial than the gender differences at the mean, which may suggest the need to rethink education policies for low-achieving boys in reading and for high-achieving girls in mathematics and science.
- Variance
- The “greater male variability hypothesis” is confirmed, although our results suggest that this may depend on the test content, the study design and the computed statistics.

To go further

- Systematically explore the effect of the test composition in terms of format of questions

Thank you

ariane.baye@ulg.ac.be