ASSESSMENT

Global Sex Differences in Test Score Variability

International testing results show greater variance in boys' scores than in girls' scores.

Stephen Machin[1] and Tuomas Pekkarinen[2]

Do boys and girls differ in their intellectual and cognitive abilities and, if so, in what way? These questions have raised considerable debate, both in terms of average performance and in terms of variability around the average.

Empirical research on gender differences in achievements produces mixed conclusions, with some evidence that favors boys and some that favors girls (*1*–*4*). In many countries, girls show superior performance in school examinations, which is also reflected in higher rates of attendance in tertiary education. In addition, girls have been improving their position relative to boys (*5*, *6*). In countries with a more gender-equal culture, the gender gap that is usually in favor of boys in average mathematics test scores is erased or even reversed in favor of girls (*7*).

At the same time, some research focuses on the notion that there are more males at the upper end of the distributions of educational and professional success (*8*). Oft-cited examples include there being more male than female Nobel Prize winners and the inequity of wages in the labor market in favor of males (*9*, *10*). Studies of talented individuals who succeed at the very highest levels, especially in science, highlight substantial male overrepresentation (*11*).

These outcomes can be generated by various kinds of distributions describing the educational and intellectual make-up of boys and girls. Differences in the gender composition of the high-scoring group can be a consequence of gender differences in the mean or variance of the test scores or both. Given that recent research has shown that gender differences in the mean test scores are not a robust phenomenon internationally (*7*), it is of interest to explore whether the gender differences in the variance of test scores are a more universal outcome.

The idea that males are intellectually and educationally more variable than females dates back a long time and is embedded in cultural history. Early academic entries are represented by late 19th century texts from Ellis (*12*) and Galton (*13*). Modern comment [e.g., (*14*)] can be quite critical of the motives of those who postulate that gender differences in variances exist.

Most of the evidence regarding gender-dependent differences in the mean and variance of intellectual and cognitive abilities is based on U.S. data (*15*, *16*), because most of the data containing test scores are from the United States. But data sources from other countries exist as well. We have therefore investigated whether the phenomenon of "higher variance" is an accurate characterization of boys' educational performance relative to girls from a wider sample of countries (*17*).

GENDER GAPS VARY BUT DIFFERENCES IN VARIABILITY ARE A ROBUST PHENOMENON

Reading	Global	OECD	Non-OECD	United States
Mean M-F gap	-0.30**	-0.32**	-0.27**	-0.32**
M-F gap at 5th quantile	-0.45**	-0.53**	-0.41**	-0.46**
M-F gap at 95th quantile	-0.20**	-0.23**	-0.14**	-0.20**
M-F Variance Ratio	1.15**	1.17**	1.19**	1.17**
M-F Ratio in bottom 5%	2.20**	2.38**	2.05**	2.45**
M-F Ratio in top 5%	0.58**	0.56**	0.68**	0.60**
Mathematics	**Global**	**OECD**	**Non-OECD**	**United States**
Mean M-F gap	0.10**	0.10**	0.11**	0.07*
M-F gap at 5th quantile	0.03**	0.02	0.06**	-0.11**
M-F gap at 95th quantile	0.20**	0.22**	0.22**	0.22**
M-F Variance Ratio	1.13**	1.13**	1.16**	1.19**
M-F Ratio in bottom 5%	0.94**	0.98	0.86**	1.21**
M-F Ratio in top 5%	1.70**	1.69**	1.70**	1.72**

Gaps and ratios. In the table, M refers to male and F to female. Gaps are M-F differences at the mean, 5th, and 95th quantiles in standard deviation units. Variance ratio refers to male variance divided by female variance, and ratios in top and bottom 5% refer to the ratio of number of males and females who scored in the top and bottom 5% of the distribution. Results for the full data are presented here; USA is excluded from OECD results. For full results, see SOM, *$P < 0.05$; **$P < 0.01$.

Research Method and Findings

We examined boy-girl differences in test score performance from the 2003 OECD (Organization for Economic Co-operation and Develop-ment) Programme for Inter-national Student Assessment (PISA), a survey of 15-year-olds who are enrolled in full- or part-time education in industrialized countries (*18*). The PISA survey contains data on individual student performance for 41 countries. Here, we analyze mathematics and reading test scores by country, focusing on differences in the mean and variance of these test scores. The test scores are standardized to have a mean of 500 and a standard deviation of 100 across the OECD countries.

Reading. We calculated the international standardized mean differences in reading test scores as the sex difference in means divided by the standard deviation, and differences in variance as the male-to-female variance ratio. In all 41 countries, the boy-girl mean difference is negative, indicating that girls generally outscore boys (see table, top, and table S1). In 35 out of 41 countries, the boy-girl variance ratio indicates, with statistical significance ($P < 0.05$), that the population of boys displays greater variance than does the population of girls (see the table above and table S1).

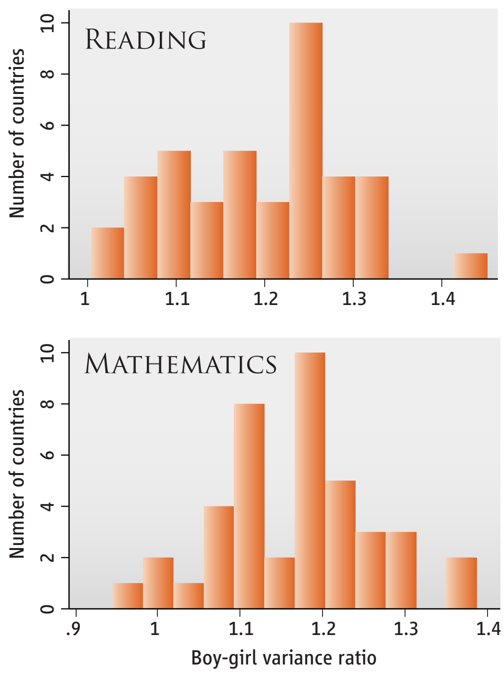
[1]Department of Economics, University College London, London, WC1 6BT, UK, and Centre for Economic Performance, London School of Economics, London, WC2A 2AE, UK. [2]Department of Economics, Helsinki School of Economics, 00101 Helsinki, Finland; Government Institute for Economic Research (VATT), 00101 Helsinki, Finland; and Department of Econonmics and Statistics, Åbo Akademi University, 20500 Turku, Finland.

*Author for correspondence. E-mail: tuomas.pekkarinen@hse.fi

Mathematics. With the same comparisons for mathematics, the mean difference is positive, which indicates that boys generally outscore girls (see table, bottom, and table S2). In 37 of 41 countries, the boy-girl variance ratio indicates with statistical significance ($P < 0.05$) that the population of boys displays greater variance than does the population of girls (see the figure below and table S2).

Tip and tail. There are differences for the two tests. On both math and reading tests, boys predominate in two of the four extreme scoring categories (low reading, high math), with girls predominating in the high reading and low math category (see table, top and bottom, and columns 4 and 5 of tables S1 and S2). For mathematics, in 35 of the 41 countries, there are more boys than girls in the top 5%. For reading, 36 of 41 countries have more girls than boys in the top 5% of scores, and 39 of 41 countries have more boys than girls in the bottom 5% of scores.

Gender gaps. Gender gaps at the 5th and 95th quantiles of the reading distribution favor girls, but tend to be smaller at the 95th quantile than at the 5th quantile. In mathematics, on the other hand, the gender difference at the 5th quantile is statistically significant at the 5% level only in 18 countries, and the sign of this difference varies across countries. In mathematics, the gender difference at the 95th quantile favors boys in 36 countries ($P < 0.05$) (see table S2).



Variance ratios in countries that participated in PISA 2003. In the majority of countries, male variability is higher than female variability in both reading and math.

Variance and test performance. For both reading and mathematics tests in all 41 countries, the boy-girl variance ratio shows a positive correlation ($r = 0.21$ for reading and $r = 0.26$ for mathematics) with the mean test score performance (fig. S1, A and B). In countries with better test score performance, the boy-girl variance ratio is significantly higher ($P < 0.05$) than in countries where the children score more poorly (table S3). We find no relation between the Gender Gap Index of the World Economic Forum (*19*) and the variance ratios for either math or reading (table S3 and fig. S2, A and B), unlike the relation previously identified between the Gender Gap Index and the gender gap in average test scores (*7*).

Other data sources. To further probe the robustness and wider applicability of these findings, we have also studied gender differences in variance from other international data sources. We analyzed results from (i) the Trends in International Mathematics and Science Study (TIMMS) given in 2003, which tested 9- and 13-year-olds in mathematics and science; and (ii) the Progress in International Reading Literacy Study (PIRLS), given in 2001, which tests reading skills of 10-year-olds (*20*, *21*). Data from both studies indicate higher variance for boys in most countries (table S4).

Implications and Conclusions

Our analysis of international test score data shows a higher variance in boys' than girls' results on mathematics and reading tests in most OECD countries. How this translates into educational achievement is a matter open for discussion. Higher variability among boys is a salient feature of reading and mathematics test performance across the world. In almost all comparisons, the age 15 boy-girl variance difference in test scores is present. This difference in variance is higher in countries that have higher levels of test score performance.

Sex differences in means are easier to characterize: It is evident from the PISA data that boys do better in mathematics, and girls do better in reading. This has a compositional effect on the variance differences as well. The higher boy-girl variance ratio in mathematics comes about because of an increased prevalence of boys in the upper part of the distribution, but the higher variance in reading is due to a greater preponderance of boys in the bottom part of the test score distribution. Because literacy and numeracy skills have been shown to be important determinants of later success in life (for instance, in terms of earning higher wages or getting better jobs), these differing variances have important economic and social implications (*22*).

We therefore confirm that 15-year-old boys do show more variability than girls in educational performance, with specifics that differ according to whether mathematics or reading are being studied and tested. These results imply that gender differences in the variance of test scores are an international phenomenon and that they emerge in different institutional settings.

References and Notes

1. A. Jensen, in *Intelligence: Genetic and Environmental Influences*, R. Cancro, Ed. (Grune and Stratton, New York, 1971).
2. E. Maccoby, C. Jacklin, *The Psychology of Sex Differences* (Stanford Univ. Press, Stanford, 1974)
3. J. S. Hyde, *Am. Psychol.* **60**, 581 (2005).
4. D. Kimura, *Neuroendocrinol. Lett.* **23**, 67 (2002).
5. S. Machin, S. McNally, *Oxf. Rev. Econ. Policy* **21**, 357 (2005).
6. J. S. Hyde, S. M. Lindberg, M. C. Linn, A. B. Ellis, C. C. Williams, *Science* **321**, 494 (2008).
7. L. Guiso, F. Monte, P. Sapienza, L. Zingales, *Science* **320**, 1164 (2008).
8. K. Bayard, J. Hellerstein, D. Neumark, K. Troske, *J. Labor Econ.* **21**, 887 (2003).
9. S. Bertsch McGrayne, *Nobel Prize Women in Science: Their Lives, Struggles, and Momentous Discoveries* (Joseph Henry Press, National Academy of Sciences, Washington, DC, 2001).
10. W. Arulampalam, A. L. Booth, M. L. Bryan, *Ind. Labor Relat. Rev.* **60**, 163 (2007).
11. C. Muller *et al.*, *Science* **307**, 1043 (2005).
12. H. Ellis, *Man and Woman: A Study of Secondary and Tertiary Sexual Characters* (Heinemann, London, 1894).
13. F. Galton, *Hereditary Genius: An Inquiry into Its Laws and Consequences* (Macmillan, London, 1869).
14. N. Noddings, *Rev. Educ. Res.* **62**, 85 (1992).
15. A. Feingold, *Rev. Educ. Res.* **59**, 185 (1992).
16. L. V. Hedges, A. Nowell, *Science* **269**, 41 (1995).
17. Materials and methods are available as supporting material on *Science* Online.
18. *PISA, The 2003 Assessment Framework – Mathematics, Reading, Science, and Problem Solving Knowledge and Skills* (OECD, Paris, 2003); www.pisa.oecd.org/.
19. R. Hausmann, L. D. Tyson, S. Zahidi, *The Global Gender Gap Report* (World Economic Forum, Geneva, Switzerland, 2006).
20. M. O. Martin, Ed., *TIMMS 2003 User Guide for the International Database* (TIMSS & PIRLS International Study Center, Boston College, Chestnut Hill, MA, 2005); http://timss.bc.edu/.
21. E. J. Gonzales, A. M. Kennedy, Eds., *PIRLS 2001 User Guide for the International Database* (Boston College, Chestnut Hill, MA, 2003); http://timss.bc.edu/.
22. E. Leuven, H. Oosterbeek, H. Van Ophen, *Econ. J.* **114**, 466 (2004).

Supporting Online Material
www.sciencemag.org/cgi/content/full/322/5906/1331/DC1

10.1126/science.1162573