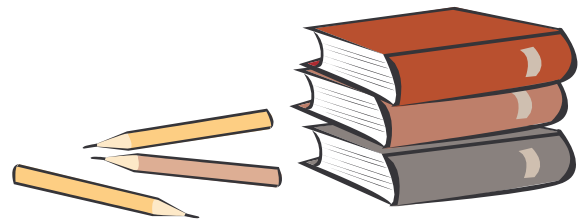


DIVERSITY

Gender Similarities in Mathematics and Science

Janet Shibley Hyde^{1*} and Marcia C. Linn²



Boys and girls have similar psychological traits and cognitive abilities; thus, a focus on factors other than gender is needed to help girls persist in mathematical and scientific career tracks.

The role of gender in mathematics and science education is hotly contested (1). Some advocate single-sex schooling as a way to let girls' talents blossom, freed from boys' domination in the classroom and from sexual harassment by boys (2, 3).

Enhanced online at www.sciencemag.org/cgi/content/full/314/5799/599

Others emphasize gender differences in learning styles, citing evidence that

girls perform best under cooperative learning conditions and boys perform best in competitive learning environments (4). These arguments rest on the assumption that psychological gender differences are large and exist in numerous domains. Should gender be the major factor in deciding which school a child attends or which curriculum a child receives? The Gender Similarities Hypothesis (5) provides an alternative view, stating instead that males and females are very similar on most, but not all, psychological variables.

Evidence for Gender Similarities

A review of meta-analyses of research on psychological gender differences identified 46 reports, addressing a variety of psychological characteristics, including mathematical, verbal, and spatial abilities; aggression; leadership effectiveness; self-esteem; and computer use (5). These research syntheses summarize more than 5000 individual studies, based on the testing of approximately 7 million people. The findings are represented on a common scale based on the standard deviation. The magnitude of each gender difference was measured using the *d* statistic (6), $d = (M_M - M_F)/s_w$, where M_M is the mean score for males, M_F is the mean score for females, and s_w is the pooled within-sex standard deviation. The *d* statistic measures the

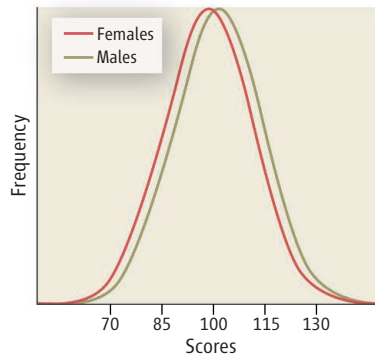
distance between male and female means, in standard deviation units. In each individual meta-analysis, the values of *d* from multiple investigations of the same outcome were weighted by sample size and combined.

A total of 124 synthesized effect sizes resulting from meta-analysis were extracted from the reports. Following convention, *d* values in the range 0.11 to 0.35 were classified as small, 0.36 to 0.65 as moderate, and 0.66 to 1.00 as large (6). Values greater than 1.0 were categorized as very large and values between 0 and 0.10 were considered trivial.

Of the effects for gender differences, 30% were trivial and an additional 48% were small. That is, 78% of the effects for psychological gender differences were small or near zero. For example, for mathematics problem-solving, $d = 0.08$ (7); for leadership effectiveness, $d = -0.02$ (8); and for negotiator competitiveness, $d = 0.07$ (9).

An essential implication of these findings is that the overlap of distributions for males and females is substantial for most outcomes. For example, the chart above shows the distribution of male and female performance for a small effect size of 0.20. Assuming that performance fits a normal distribution (supporting online material text), for means that are 0.20 standard deviations apart ($d = 0.20$), the populations show 85.3% overlap. Only 54% of members of one gender exceed the 50th percentile for the other gender. For $d = 0.10$, there is 92.3% overlap in the distributions of the two groups, and 52% of those of one gender exceed the 50th percentile for the other. Even for a moderate effect size ($d = 0.50$), 60% of members of one gender exceed the 50th percentile for the other gender.

Most relevant here is the meta-analysis of research on gender differences in mathematics



Effect size of 0.20. When the effect size between two groups is 0.20, 85.3% of the distributions overlap.

performance, which was based on 100 studies and the testing of more than 3 million people (7). Patterns emerged as a function of the age of test takers and the cognitive level of the test. Girls outperformed boys on computation in elementary school and middle school ($d = -0.20$). There was no gender difference in high school. There was no gender difference in deeper understanding of mathematical concepts at any age.

For complex problem solving, a skill that is highly relevant for science, technology, engineering, and mathematics careers, there was no gender difference in elementary or middle school; a small difference favoring boys emerged in high school ($d = 0.29$). Consistent with these findings of gender similarities in mathematics performance, in 2001 women earned 48% of the bachelor's degrees in mathematics in the United States (10), demonstrating that substantial numbers of women do have the ability to engage mathematics successfully at the advanced levels required of a mathematics major.

A few exceptions to the pattern of gender similarities were identified in the review of meta-analyses. Most relevant to educational settings are the gender differences in activity level and physical aggression, with effect sizes for aggression ranging between 0.40 and 0.60 (males more aggressive) across several meta-analyses (11–14). Males display a higher activity level than females, $d = 0.49$ (15). It is widely believed that there are substantial gender differences in interests in fields such as psychology compared with physics, although we found no meta-analysis of this research literature.

Gender Similarities in Science

The National Assessment of Educational Progress (NAEP), the Nation's Report Card, tests thousands of students across the United States on achievement in mathematics and

¹Department of Psychology, University of Wisconsin, Madison, WI 53706, USA. ²Education in Mathematics, Science, and Technology, University of California, Berkeley, CA 94720, USA.

*To whom correspondence should be addressed. E-mail: jshyde@wisc.edu

science. The conclusions from 2005 stated that males outperformed females at all three grades tested in science (16).

Another look at the data leads to a different conclusion. For fourth-graders, the average science score was 152.53 (SD = 32) for boys and 148.66 (SD = 30) for girls. That is, the difference is less than 4 points on a scale that ranges from 0 to 300. As NAEP reports, this gender difference is statistically significant, given the large sample size (roughly 100,000 students per grade). However, the effect size, d , for this gender difference is 0.12, reflecting a small difference. Increasingly large samples can detect increasingly small differences, but an assessment of the effect size, d , gives a more accurate reflection of the importance of the difference. Emphasis on statistical significance, while ignoring the magnitude of the effect, risks exaggerating the importance of the differences (17).

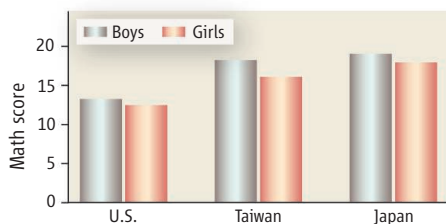
The magnitude of the gender gap in science achievement remains similar from fourth to 12th grade. NAEP reports average scores of 149.01 for boys (SD = 35) and 145.15 for girls (SD = 33) in 12th grade, for an effect size of $d = 0.11$, a small difference that is no larger than the difference in fourth grade. The NAEP data provide better evidence for gender similarities in science achievement than they do for gender differences.

Cross-National Comparisons

Cross-national data provide another way to evaluate the magnitude of gender differences in mathematics and science performance. Research on fifth-graders' performance on mathematical word problems in the United States, Taiwan, and Japan reveals a slight male advantage in each culture, but a larger difference between cultures (see chart above) (18). Whereas the effect size for the gender difference in the United States is $d = 0.18$, the effect size for the difference between U.S. boys and Japanese boys is $d = 1.42$. These cultural differences reflect many factors, including differences in curriculum, time spent in homework, and parents' beliefs in the importance of effort in school performance (19). Even within the United States, the gender difference in mathematics performance is larger in some ethnic groups than others. For whites, $d = 0.13$; whereas for blacks, $d = -0.02$; for Hispanics, $d = 0.00$; and for Asian Americans, $d = -0.09$ (7).

Implications

Women earn 46% of the Ph.D.'s in biology but, despite evidence for gender similarities, they earn only 25% of the Ph.D.'s in physical science and 15% in engineering. Women comprise 30% of the assistant professors in biology but



Cross-national and gender differences in math.

Differences in fifth-graders' performance on word problems are larger between countries than between genders (18). Boys' scores are shown in blue; girls' scores are shown in red.

only 16% in physical science and 17% in engineering (20). Too often, small differences in performance in the NAEP and other studies receive extensive publicity, reinforcing subtle, persistent, biases (20, 21). Indeed, the magnitude of the attitudinal association between science and males is large, $d = 0.72$ (22, 23). These biases can have an impact on decisions about admissions, hiring, and promotion (24, 25). These biases may contribute to popular beliefs about same-sex education and learning styles, and dissuade some individuals from persisting in science (26).

For example, advocates of same-sex education claim advantages for both boys and girls (2). Some argue that boys' great activity level and aggressiveness make it difficult for girls to learn and participate actively, and at the same time, boys need a classroom that tolerates their active style. Activity level and physical aggression are two exceptions to the gender similarities rule, with effect sizes around 0.50 for each. Yet even a gender difference of that magnitude means that 40% of one group (in this case, girls) score higher than the average for the other group (boys). If the idea is to separate children into classrooms for the more active and aggressive and the less active and aggressive, gender is not the best indicator. A teacher's rating of activity level would be far more accurate.

The phenomenon of gender similarities has implications for schooling. Emphasis on gender differences in the popular literature reinforces stereotypes that girls lack mathematical and scientific aptitude. However, gender is a poor indicator of whether one will major in mathematics or the biological sciences as an undergraduate. A better predictor would be actual mathematics achievement scores in middle school or high school (27). A cultural overemphasis on gender differences may mask critical predictive variables and lead to decision-making that is empirically unsupported. To help teachers succeed, we may need to address variability in aggression and activity level for all learners. To neutralize traditional stereotypes about girls' lack of

ability and interest in mathematics and science, we need to increase awareness of gender similarities. Such awareness will help mentors and advisers avoid discouraging girls from entering these fields. Continued monitoring of the relative progress of boys and girls is essential so that neither group falls behind.

Rather than focusing on gender differences, mathematics and science educators and researchers could more profitably examine ways to increase awareness of the similarities in performance and in ability to succeed.

References and Notes

1. S. Ceraulo, *J. Chem. Educ.* **76**, 615 (1999).
2. J. L. Streitmatter, *For Girls Only: Making a Case for Single-Sex Schooling* (SUNY Press, Albany, NY, 1999).
3. M. Warrington, M. Younger, *Oxford Rev. Educ.* **27**, 339 (2001).
4. M. Gurian, A. Ballew, *The Boys and Girls Learn Differently Action Guide for Teachers* (Jossey-Bass, San Francisco, CA, 2003).
5. J. S. Hyde, *Am. Psych.* **60**, 581 (2005).
6. J. Cohen, *Statistical Power Analysis for the Behavioral Sciences* (Erlbaum, Hillsdale, NJ, ed. 2, 1988).
7. J. S. Hyde, E. Fennema, S. Lamon, *Psych. Bull.* **107**, 139 (1990).
8. A. H. Eagly, S. Karau, M. Makhijani, *Psych. Bull.* **117**, 125 (1995).
9. A. E. Walters, A. Stuhlmacher, L. Meyer, *Org. Behav. Hum. Dec. Proc.* **76**, 1 (1998).
10. National Science Foundation, *Women, Minorities, and Persons with Disabilities in Science and Engineering* (NSF, Arlington, VA, 2004); available online (www.nsf.gov/statistics/wmpd/underdeg.htm).
11. J. S. Hyde, *Dev. Psych.* **20**, 722 (1984).
12. J. S. Hyde, M. C. Linn, *The Psychology of Gender: Advances Through Meta-Analysis* (Johns Hopkins Univ. Press, Baltimore, MD, 1986).
13. A. H. Eagly, V. Steffen, *Psych. Bull.* **100**, 309 (1986).
14. J. Archer, *Rev. Gen. Psychol.* **8**, 291 (2004).
15. W. O. Eaton, L. R. Enns, *Psych. Bull.* **100**, 19 (1986).
16. The Nation's Report Card (http://nationsreportcard.gov/science_2005).
17. L. Wilkinson, *Am. Psych.* **54**, 594 (1999).
18. M. Lummis, H. W. Stevenson, *Dev. Psych.* **26**, 254 (1990).
19. H. W. Stevenson, C. Chen, S.-Y. Lee, *Science* **259**, 53 (1993).
20. J. Handelsman et al., *Science* **309**, 1190 (2005).
21. J. A. Bargh, T. L. Chartrand, *Am. Psych.* **54**, 462 (1999).
22. B. Nosek, M. Banaji, A. Greenwald, *Group Dyn. Theory Res., Pract.*, **6** 101 (2002).
23. B. Nosek, M. Banaji, A. Greenwald, *J. Pers. Soc. Psych.* **83**, 44 (2002).
24. B. Bergmann, *In Defense of Affirmative Action* (BasicBooks, New York, 1996).
25. M. C. Linn, *Educ. Res.* **27**, 15 (1998).
26. E. Seymour, N. Hewitt, *Talking About Leaving* (Westview Press, Boulder, CO, 1997).
27. R. H. Tai, C. Q. Liu, A. V. Maltese, X. Fan, *Science* **312**, 1143 (2006).
28. This study was funded in part through grants from NSF, REC 0635444, 0207109, 9980620, and 0233649, and ESI-0242701, 9720384, and 0334199. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Supporting Online Material

www.sciencemag.org/cgi/content/full/314/5799/599/DC1

10.1126/science.1132154