Gender, culture, and mathematics performance

Janet S. Hyde[a,1] **and Janet E. Mertz**[b]

[a]*Department of Psychology, University of Wisconsin, Madison, WI 53706; and* [b]*McArdle Laboratory for Cancer Research, University of Wisconsin School of Medicine and Public Health, Madison, WI 53706-1599*

Edited by Randy Schekman, University of California, Berkeley, CA, and approved April 2, 2009 (received for review February 5, 2009)

Using contemporary data from the U.S. and other nations, we address 3 questions: Do gender differences in mathematics performance exist in the general population? Do gender differences exist among the mathematically talented? Do females exist who possess profound mathematical talent? In regard to the first question, contemporary data indicate that girls in the U.S. have reached parity with boys in mathematics performance, a pattern that is found in some other nations as well. Focusing on the second question, studies find more males than females scoring above the 95th or 99th percentile, but this gender gap has significantly narrowed over time in the U.S. and is not found among some ethnic groups and in some nations. Furthermore, data from several studies indicate that greater male variability with respect to mathematics is not ubiquitous. Rather, its presence correlates with several measures of gender inequality. Thus, it is largely an artifact of changeable sociocultural factors, not immutable, innate biological differences between the sexes. Responding to the third question, we document the existence of females who possess profound mathematical talent. Finally, we review mounting evidence that both the magnitude of mean math gender differences and the frequency of identification of gifted and profoundly gifted females significantly correlate with sociocultural factors, including measures of gender equality across nations.

exceptional talent | gender gap index | greater male variability hypothesis | International Mathematical Olympiad | Programme for International Student Assessment

Researchers first began investigating gender differences in abilities and behaviors in the 1880s (1). The scientists of the time concluded that women's smaller brains were sadly deficient. For example, George Romanes declared in 1887 that mental abilities were secondary sex characteristics attributable to brain size (2). Twenty-first century scientists have vastly better research methods available to them. Moreover, the behaviors and performance of women and men in 2009 are substantially different from what they were in the Victorian era. This article reviews and synthesizes the current evidence on gender differences in abilities, focusing on mathematical skills because of the crucial role they play in success in careers in science, technology, engineering, and mathematics, i.e., STEM fields. The review is organized around 3 questions: Do gender differences in mathematics performance exist in the general population? Do gender differences exist among the highly mathematically talented? Do females exist who possess profound mathematical talent? Last, we consider the evidence concerning the contribution of sociocultural factors to the gender differences observed in measured mathematical performance.

Do Gender Differences in Mathematics Performance Exist in the General Population?

In influential reviews published in 1966 and 1974, the noted developmental psychologist Eleanor Maccoby concluded that gender differences in mathematics performance were scientifically well established, with males scoring higher (3, 4). She documented that boys and girls acquire early number concepts similarly in the preschool years, a conclusion fully supported by contemporary data (5), and that their performance throughout elementary school was similar; however, boys' skills in mathematics increased faster than girls' beginning around 12 or 13 years of age, creating a significant gender gap in performance by high school.

The technique of meta-analysis became available by the 1980s. It provides a powerful statistical method for synthesizing the results of numerous studies on a given question. In research on gender differences, the meta-analyst computes the effect size, d, for each study and then computes a weighted average effect size across all studies (6). The effect size is computed as $d = (M_M - M_F)/S_w$, where M_M is the mean score for males (M), M_F is the mean score for females (F), and S_w is the within-groups standard deviation. Thus, it is a measure of the distance between the male and female means in standard deviation units. Positive values represent better performance by males, whereas negative values represent better performance by females. According to standard guidelines, an effect size of 0.20 is small, 0.50 is moderate, and 0.80 is large (7).

Hyde and colleagues reported a 1990 meta-analysis on gender differences in mathematics performance involving 100 studies representing the testing of >3 million individuals, most from the U.S. but some from other nations such as Australia and Canada (8). Overall, they found $d = -0.05$ for samples of the general population, an effect so small as to be considered no gender difference. Further analyses explored effects of age and cognitive level of test items on the magnitude of gender difference. Test items were coded as assessing simple computation (i.e., memorized math facts), deeper understanding of concepts, or, at the highest level, complex problem solving. The results indicated a slight female advantage in computation in elementary and middle school, and no difference in high school. There were no gender differences in understanding of concepts at any age. Complex problem solving displayed no gender difference in elementary school and middle school, but a gender difference favoring males emerged in high school, with $d = 0.29$. This latter finding is of concern because complex problem solving is an essential skill for success in life and in STEM careers.

These findings were largely replicated in a 1995 meta-analysis using large datasets based on the testing of excellent probability samples of U.S. adolescents (9). For high school students, d values ranged between 0.03 and 0.26 for mathematics performance, that is, boys performed better than girls by a small amount.

One prominent explanation for this measured gender difference in math performance in high school has been differential patterns of course taking

Author contributions: J.S.H. and J.E.M. analyzed data and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: jshyde@wisc.edu.

Table 1. Gender differences in mathematics performance in U.S. state assessments

	d	VR
Grade 2	0.06	1.11
Grade 3	0.04	1.11
Grade 4	−0.01	1.11
Grade 5	−0.01	1.14
Grade 6	−0.01	1.14
Grade 7	−0.02	1.16
Grade 8	−0.02	1.21
Grade 9	−0.01	1.14
Grade 10	0.04	1.18
Grade 11	0.06	1.17

d, mean score for males minus mean score for females divided by the pooled within-gender standard deviation; VR, variance ratio.

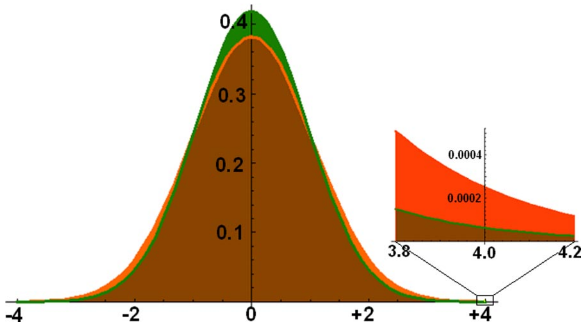

Fig. 1. Theoretical normal distributions for males (orange line) and females (green line) when their means are identical and the M:F VR = 1.2. The schematic on the right shows a blowup of the distributions in the region from 3.8 to 4.2 standard deviations above the mean. Brown, area of overlap of the 2 distributions; green and orange, areas unique to females and males, respectively.

(10, 11), that is, girls were less likely than boys to take advanced mathematics courses in high school. They were also less likely to take chemistry and physics, other courses where complex problem solving is taught. Lacking this training, girls, not unexpectedly, performed less well than boys on standardized tests.

However, gender patterns had changed by the beginning of the 21st century. Girls are now taking calculus in high school at the same rate as boys, although they still lag behind boys in taking physics (12). In this new environment, do boys' and girls' math scores still differ? Massive amounts of data relating to this question are available because No Child Left Behind (NCLB) legislation in the U.S. mandates that all states test all children in all grades on their proficiency in mathematics. In one recent study, researchers obtained useable data from 10 states representing the testing of >7 million youth (13). Averaged across these 10 states, gender differences in performance were close to zero in all grades, including high school, with d values ranging between −0.02 and 0.06 (Table 1). When analyzed by ethnicity, the same pattern of gender similarities was found for all ethnic groups studied, that is, African Americans, Latinos, Asian Americans, American Indians, and Whites. Thus, girls have now reached parity with boys in mathematics performance in the U.S., even in high school where a gap existed in earlier decades.

However, coding of the test items on these examinations for cognitive level indicated that none of them tapped complex problem solving at most grade levels for most states (13). Thus, it was impossible with these NCLB datasets to investigate whether a gender gap existed in complex problem solving. Therefore, the researchers also examined data from the National Assessment of Educational Progress (NAEP), a federally managed program that tests a random sample of U.S. students each year (14). Items from 12th grade data categorized by NAEP as hard and by the researchers as requiring complex problem solving were analyzed for gender differences; effect sizes were found to average $d = 0.07$, a trivial difference. These findings provide further evidence that U.S. girls have now reached parity with boys, even in high school, and even for measures requiring complex problem solving.

Some have argued that the absence of gender differences in mathematics performance in the general population is irrelevant to the advancement of STEM fields; rather, researchers should focus on the mathematically talented, a topic discussed below. However, Weinberger found that <1/3 of the college-educated white U.S. males in the STEM workforce had high school quantitative SAT scores >650 (15). Thus, progress in STEM fields is fueled, not only by the highly talented, but also by the millions of laboratory technicians and other bachelors- and masters-level scientists whose mathematics skills might place them below the 75th percentile, but whose contributions are still essential.

Moreover, numeracy is important for everyone, with mathematical competency being crucial to anyone shopping for a home mortgage, investing their savings for retirement, or deciding among several treatment options for a serious medical ailment. The recent example of consumers' failure to comprehend adjustable-rate mortgages is a sobering case in point. Mathematical skills are essential, not only for accountants, economists, and physicists, but also for teachers, nurses, politicians, and the lay public in general.

Do Gender Differences Exist Among the Mathematically Talented?

The hypothesis that the variability of intellectual abilities is greater among males than females was originally proposed by Ellis in 1894 to explain a phenomenon that seemed obvious at the time: There were both an excess of males among the mentally defective and very few female geniuses (1). If this Greater Male Variability Hypothesis were valid, it could account for the existence of a preponderance of males at the highest levels of performance even when a mean gender difference does not exist, as shown schematically in Fig. 1. This is the hypothesis to which Lawrence Summers was referring when he stated at the National Bureau of Economic Research Conference held on January 14, 2005, "There are issues of intrinsic aptitude, and particularly of the variability of aptitude, and that those considerations are reinforced by what are in fact lesser factors involving socialization and continuing discrimination. It's talking about people who are 3 ½, 4 standard deviations above the mean in the one-in-5,000, one-in-10,000 class. Even small differences in the standard deviation will translate into very large differences in the available pool substantially out."

The statistic used nowadays to test this hypothesis is called the variance ratio (VR), that is, the ratio of male variance to female variance in a distribution. Thus, variance ratios >1.00 indicate greater male variability. Variance ratios calculated from the state math assessments (13) are shown in Table 1. All are >1.00, but the discrepancy in variances from gender similarity is not great, with VRs ranging between 1.11 and 1.21.

Theoretical models have been used to examine the consequences of greater male variance based on the assumption that the scores are normally distributed (16). For example, if $d = 0.05$ and $VR = 1.12$, values representative of the ones found in the state assessments, then the ratio of males-to-females scoring above the 95th percentile would be 1.34. At a very high cutoff, the 99.9th percentile, the M:F ratio would be 2.15. Even for a STEM specialty that requires mathematics skills at the latter level, we would expect workers in the occupation to be 68% male and 32% female if math talent were the only factor that mattered. Yet in recent years, for example, women accounted for only 18% of the Engineering Ph.D.s awarded in the U.S. (17). If $d = 0.00$ and VR = 1.20 as shown in Fig. 1, a field would need to require workers who are at least 4 standard deviations (SDs) above the mean, the 1-in-20,000 level, to be only 18% females and 5 SDs, the 1-in-one million level, to be only 9% females.

Theoretical models are, of course, just that. Actual distributions rarely conform exactly to normal ones. Thus, gender ratios in the upper tails of actual distributions were calculated using data from the Minnesota state assessments (13). Results were analyzed separately by ethnicity to ensure that the findings were not limited to the predominantly White samples that have been the mainstay of U.S. research. For students scoring above the 95th percentile, the M:F ratio was 1.45 for Whites, close to theoretical prediction. At the 99th percentile, the M:F ratio was 2.06, again close to theoretical prediction. However, the M:F ratio was only 0.91 for Asian-Americans, that is, more girls than boys scored above the 99th percentile. Analysis of data from 15-year-old students participating in the 2003 Program for International Student Assessment (PISA) likewise indicated that as many, if not more girls than boys scored above the 99th percentile in Iceland, Thailand, and the United Kingdom (18). The M:F ratios above the 95th percentile on this examination also fell between 0.9 and 1.1 for these above-named countries plus Indonesia, that is, were not significantly different from equal variances (19). These findings challenge the Greater Male Variability Hypothesis, which, if valid, should hold for all representative populations, regardless of ethnicity or nationality.

Two recent studies directly address the question of whether greater male variability in mathematics is a ubiquitous phenomenon. Machin and Pekkarinen (19) reported that the M:F VR in mathematics was significantly >1.00 at the $P < 0.05$ level among 15-year-old students in 34 of 40 countries participating in the 2003 PISA and among 13-year-old students in 33 of 50 countries participating in the 2003 Trends in International Mathematics and Science Study (TIMSS). However, these data also indicated that the math VR was significantly less than or insignificantly different from 1.00 for some of the countries that participated in these assessments (e.g., Table 2), a finding inconsistent with the Greater Male Variability Hypothesis.

Table 2. Differences in variability in math performance between boys and girls among some selected nations

Country	2003 PISA 15 year olds, M/F VR†	1995 TIMSS 17 year olds $(SD_M - SD_F)/SD_W$‡
Canada	1.24*	0.05
Czech Rep.	1.07	0.11
Denmark	0.99	0.01
Germany	1.12*	−0.05
Iceland	1.24*	0.04
Indonesia	0.95*	ND
Ireland	1.07	ND
Lithuania	ND	−0.06
Mexico	1.08*	ND
Netherlands	1.00	−0.13
Thailand	1.10*	ND
Tunisia	1.03	ND
Russian Fed.	1.20*	0.02
Slovenia	ND	0.01
Switzerland	1.11*	0.02
UK	1.06*	ND
USA	1.19*	0.09

*, VR significantly different from 1.0, $P < 0.05$. ND, not determined.

†Variance ratios taken from table S2 of Machin and Pekkarinen (19).

‡Calculated from data presented in table 2 of Penner (20); P values are not known.

Penner has performed a detailed analysis of the distributions of math scores obtained by boys compared with girls in each country that participated in the 1995 TIMSS (20). Striking was his finding of considerable country-to-country variation, not only in the magnitude of the difference between mean male and female scores, but also in the shapes of the distributions, ratios of males-to-females scoring in the right and left tails of the distributions, and difference in standard deviation (SD) between males and females. We have normalized these latter differences to overall within SD for each country such that the numbers >0 in the rightmost column of Table 2 indicate greater male variability. Notable is the fact that numerous countries had a normalized SD difference that was insignificantly different from zero, with 3 even having a negative value, that is, greater female variability. Neither the 10th-grade 2003 PISA nor 12th-grade 1995 TIMSS data gave any indication of greater male variability in mathematics for either Denmark or the Netherlands. As Penner concluded, "The common assumption that males have greater variance in mathematics achievement is not universally true." Given the absence of universality, the occurrence of greater male variability and scarcity of top-scoring females in many, but not all, countries and ethnic groups must be largely due to changeable sociocultural factors, not immutable, innate biological differences between the sexes.

Some studies have focused specifically on the mathematically talented. The best known example is the Study of Mathematically Precocious Youth (SMPY) or Study of Exceptional Talent (SET), an ongoing study originally begun at The Johns Hopkins University in the 1970s (21). These researchers administer the SAT to children <13 years of age who have been identified as mathematically advanced. Their sample is voluntary, and the sampling frame is not well defined. It has also changed over time with respect to sample size and ethnicity, including large numbers of children of immigrants from Eastern Europe and Asia in recent years. In 1980–1982, they reported a very lopsided M:F ratio of 13:1 among students scoring ≥700 on the quantitative section of the examination (21). However, here too, the gender gap has dramatically narrowed with time. The M:F ratio was down to 2.8:1 by 2005 (22, 23). Thus, females now represent at least 1/4 of the mathematically precocious youth being identified in this U.S. talent search. This fairly rapid and dramatic change occurred coincident with enactment of Title IX, the second wave of the women's movement, and greatly increased immigration of Eastern Europeans and Asians to the U.S., points further discussed below.

Do Females Exist Who Possess Profound Mathematical Talent?

No woman to date has been awarded a Fields Medal, the so-called "Nobel Prize of mathematics." Nevertheless, over the centuries women have made many profound contributions to mathematics, from Hypatia of Alexandria in *ca.* 400 CE to Professor Maryam Mirzakhani of Stanford University in the 21st century. Notables in between have included Marie-Sophie Germain, Ada Lovelace, Emmy Noether, Dame Mary Cartwright, Grace Hopper, and Julia Robinson (see www.maa.org/pubs/posterW.pdf for brief biographies of these and some other outstanding female mathematicians). Ingrid Daubechies, Dusa McDuff, Marina Ratner, and Karen Uhlenbeck are among the current members of the mathematics section of the U.S. National Academy of Sciences. Thus, the answer to the above question is emphatically "Yes," that is,

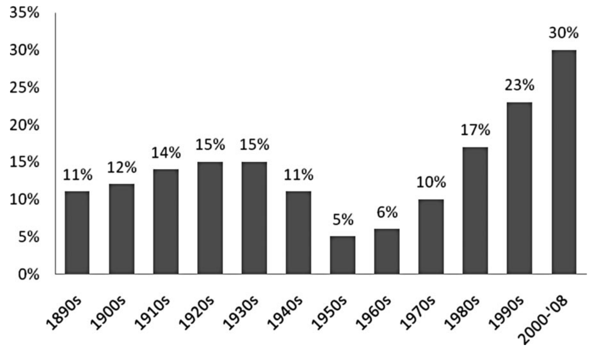

Fig. 2. Percentage of U.S. mathematics Ph.D.s awarded to women by decade. Data were taken from refs. 24–26.

females with profound mathematical talent do, in fact, exist.

With women prohibited from pursuing doctoral studies in mathematics at essentially all universities in the world before the 1890s, let alone being mathematics professors at major research universities, their extreme scarcity before the 20th century was, undoubtedly, largely due to very few women having had the opportunity to develop and use their mathematical talents. Once opportunities began to open up, women accounted for 14%–15% of the Ph.D.s awarded in mathematics in the U.S. during the decades before World War II (refs. 24 and 25 and Fig. 2). However, by the 1950s, this number had plummeted to 5% and did not return to its previous level until the 1980s. In the past decade, the percentage of U.S. mathematics Ph.D.s awarded to females among both U.S. citizens and noncitizens has been hovering at ≈30% (ref. 26 and Fig. 2), a number consistent with the M:F 2.8:1 ratio observed in recent SMPY cohorts (23) and the ≈2:1 ratios for students scoring above the 99th percentile on standardized tests (13, 18). However, U.S. research universities hired many of their STEM faculty in the 1960s post-Sputnik era when federal funding for science greatly increased and large numbers of baby boomers began to attend college, but Ph.D.-level female mathematicians were scarce. These professors have only recently begun to retire, accounting in part for the current large disparity between percentage of female professors in top-ranked research departments (27, 28) and students of mathematics.

All students identified in the SMPY or who achieve a Ph.D. in mathematics possess both a gift for mathematics and the will to study it intensively. However, only a small subset of this group is truly profoundly gifted in mathematics. To identify the latter, Andreescu et al. (28) examined data on high school and college students who excel at the very highest level in extremely difficult competitions in mathematical problem solving where solutions require the writing of rigorous proofs. One such competition is the International Mathematical Olympiad (IMO) (www.imo-official.org). The top scorers on the IMO have truly exceptional skills in mathematics, that is, at the 1-in-a-million level. Because the IMO is taken by 6-member teams consisting of very top mathematics students from ≈95 countries, it also provides information regarding cultural differences among nations.

Table 3 lists the names of some female high school students who scored among the very top in the world on the IMO. Some of them (e.g., Ana Caraiani) have also scored among the very top on the intercollegiate William Lowell Putnam Mathematical Competition (ref. 28 and www.maa.org/awards/putnam.html), an examination open only to students who have matriculated to colleges in the U.S. and Canada. Some (e.g., Maryam Mirzakhani) have become tenured professors in very top-ranked mathematics departments and won prestigious awards for their accomplishments in mathematics research. Thus, females of this profound caliber in mathematics exist.

Table 4 indicates the percentages of students on IMO teams who were female during the past 3 decades for countries whose teams achieved a median rank among the top 30 in recent years. Some of these high-ranked countries (e.g., Russia, Serbia) had >20% female team members during some decades, a number that should be considered a lower bound on the percentage of the population with profound intrinsic aptitude for mathematics who are female. Noteworthy, however, are the facts that the frequency with which females are members of IMO teams can dramatically change with time, differs quite significantly among countries, and, even, can differ 20- to 60-fold among ethnic groups within countries (Table 4 and ref. 28). For example, the U.S. had zero females on its teams throughout the first 23 IMOs in which it participated, finally having 3 females on 5 of its teams during the past 11

Table 3. Females scoring among top in world in IMO

Name	Country	Rank	Year
Zhuo Chen	People's Rep. of China	12	2008
Lisa Sauermann	Germany	12	2008
Maria Ilyukhina	Russian Federation	6	2007
Livia Alexandrra Ilie	Romania	7	2007
Sherry Gong	USA	7	2007
Maria Colombo	Italy	16	2006
Galyna Dobrovolska	Ukraine	16	2004
Ana Caraiani	Romania	6	2003
Greta Panova	Bulgaria	10	2001
Suh Hyun Choi	Rep. of Korea	15	1999
Sachiko Nakajima	Japan	7	1996
Maryam Mirzakhani	Islamic Rep. of Iran	1*	1995
Chenchang Zhu	People's Rep. of China	1*	1995
Theresia Eisenkölbl	Austria	1*	1994
Catriona Maclean	United Kingdom	1*	1994
Marianna Csörnyei	Hungary	14	1993
Eva Myers	United Kingdom	14	1992
Evgenia Malinnikova	USSR	1*,1*,11	1991, 1990, 1989
Jun Teng	People's Rep. of China	1*	1987
Olga Leonteva	USSR	5	1985
Karin Gröger	German Dem. Rep.	1*	1984
Tatyana Hovanova	USSR	2	1976
Lidia Goncarova	USSR	3	1962

IMO, International Mathematical Olympiad.
*Scored perfect 42.

Table 4. Female participants on IMO teams of top 30-ranked countries

Country	Median team rank 2000–2008	% world Population	Girls/total (% girls)		
			1978–1988	1989–1998	1999–2008
People's Rep. China	1	19.70	3/19 (15.8)	3/54 (5.6)	1/60 (1.7)
USSR/Russian Fed.	2	2.11	1/54 (1.8)	13/60 (21.7	2/60 (3.3)
USA	3	4.54	0/64 (0)	1/60 (1.7)	4/60 (6.7)
Rep. of Korea	4	0.72	0/6 (0)	3/60 (5.0)	6/60 (10.0)
Bulgaria	5	0.11	7/48 (14.6)	1/60 (1.7)	8/60 (13.3)
Vietnam	5	1.30	1/51 (2.0)	3/60 (5.0)	1/60 (1.7)
Japan	9	1.90	—	2/54 (3.7)	0/60 (0)
Taiwan	9	0.34	—	3/42 (7.1)	2/60 (3.3)
Hungary	10	0.15	2/52 (3.8)	3/60 (5.0)	5/60 (8.3)
Islamic Rep. Iran	10	1.05	0/13 (0)	3/60 (5.0)	0/60 (0)
Romania	10	0.32	0/60 (0)	0/60 (0)	7/60 (11.7)
Ukraine	13	0.69	-	1/41 (2.4)	6/60 (10.0)
GDR/Germany	15	1.22	3/47 (6.4)	4/60 (6.7)	4/60 (6.7)
West Germany	—	—	0/76 (0) for 1977–1990		
India	15	16.9	—	0/59 (0)	4/60 (6.7)
Turkey	17	1.05	ND	2/60 (3.3)	2/60 (3.3)
Belarus	18	0.14	—	1/37 (2.7)	5/60 (8.3)
Canada	19	0.50	1/48 (2.1)	6/60 (10.0)	7/60 (11.7)
Poland	19	0.57	1/58 (1.7)	1/60 (1.7)	0/60 (0)
Israel	21	0.11	0/27 (0)	0/58 (0)	3/60 (5.0)
Thailand	21	0.94	—	ND	2/60 (3.3)
United Kingdom	22	0.90	1/64 (1.6)	6/60 (10.0)	7/60 (11.7)
Yugoslavia/Serbia	23	0.15	8/64 (12.5)	6/60 (10.0)	13/60 (21.7)
Brazil	24	2.79	2/48 (4.2)	2/57 (3.5)	2/60 (3.3)
Hong Kong	24	0.11	1/6 (16.7)	2/60 (3.3)	3/60 (5.0)
Kazakhstan	24	0.23	—	0/36 (0)	1/60 (1.7)
Australia	25	0.32	2/48 (4.2)	2/60 (3.3)	5/60 (8.3)
Rep. of Moldova	26	0.06	—	5/28 (17.8)	0/58 (0)
France	30	0.96	0/64 (0)	0/60 (0)	4/60 (6.7)
Singapore	30	0.07	0/6 (0)	1/60 (1.7)	4/60 (6.7)

—, not yet participating; ND, not determined; IMO, International Mathematical Olympiad.

years. Likewise, the United Kingdom fielded only 1 female on its teams from 1967 to 1988, yet has had 10 different females on its teams during the past 2 decades, with several participating more than once. During the 13-year period immediately before reunification, the German Democratic Republic had 5 females on its teams, whereas West Germany had zero. Since partitioning, Slovakia has fielded 3 times as many females on its teams as has the Czech Republic (28). During the past decade, the Republic of Korea has had 6 female participants versus Japan's zero. Such large differences among genetically related populations and rapid changes over time within countries in the frequency of identification of females with extreme talent in mathematical problem solving cannot be primarily due to biological factors.

Role of Culture in Nurturing Mathematical Talent

Current research provides abundant evidence for the impact of sociocultural and other environmental factors on the development and nurturing of mathematical skills and talent and the size, if any, of math gender gaps. The evidence comes from both cross-ethnic and cross-national studies and the above-cited changes observed within countries over time in the general, SMPY, and IMO-level populations.

Several researchers have investigated cross-national patterns of gender differences in math performance, studies that also provide clues as to the specific cultural factors that most affect outcomes. Baker and Jones (29) found that the magnitude of the mean gender difference in mathematics performance on the Second International Mathematics Study (SIMS) significantly correlated, across nations, with measures of gender inequality. For example, the size of the math gender gap correlated −0.55 with the percentage of women in the workforce in those nations. Likewise, Guiso and colleagues (18), using 2003 PISA data testing 15-year-olds from 40 countries, found that gender inequality as measured by the World Economic Forum's Gender Gap Index (GGI) (30) significantly correlated with the magnitude of the mean math gender gap. The GGI provides a measure of the gap between men and women in economic participation and opportunity, educational attainment, political empowerment, and health and survival; the closer it is to 1.00, the smaller the gender gap on these measures. In other words, Guiso et al. (18) concluded that the math gender gap varies across nations; nations with greater gender equality typically have a smaller math gender gap.

Regarding change over time, the likeliest explanation for the dramatic improvement in math performance by U.S. females lies in 2 recent cultural trends: (*i*) girls in general taking more mathematics and science courses during high school due, in part, to changes in requirements for graduation and admission to colleges, and (*ii*) the opening up to females shortly before or after enactment of Title IX in 1972 of STEM-intensive specialty high schools, colleges, and graduate schools along with career opportunities in STEM fields. These 2 trends are intimately connected.

The increase in women pursuing careers in STEM has been quite dramatic in some STEM fields. For example, only 14% of the U.S. doctoral degrees in the biological sciences went to women in 1970, whereas this figure had risen to 49% by 2006 (31). Entry into other STEM areas has been slower, yet substantial. For example, 5.5% of U.S. doctoral degrees in the physical sciences were awarded to women in 1970, compared with 30% in 2006; the percentages in mathematics and statistics were 8% in 1970 and 32% in 2006 (26, 32). Clearly, numerous women are willing and able to learn the mathematics needed for advanced degrees in these areas when provided with an appropriately nurturing sociocultural environment along with educational and career opportunities.

The Guiso study (18) also provides data relevant to the Greater Male Variability Hypothesis discussed earlier. Whereas the U.S. ranked a dismal 36th out of 40 countries in the ratio of 15-year-old females-to-males scoring above the 99th percentile in the 2003 PISA, the United Kingdom had equal numbers of girls and boys scoring above this percentile, and Iceland and Thailand actually had more girls than boys above this cutoff. These authors concluded that a strong correlation exists between a country's measures of gender inequity and the size of the math gender gap both at the mean and the right tail of the distribution.

In new analyses, we calculate a Pearson correlation of −0.34 ($P < 0.05$) between the ratio of males-to-females scoring above the 95th percentile on the 2003 PISA (supplementary online material accompanying ref. 19) and the 2007 GGI (30). Likewise, we find a correlation of 0.44 ($P < 0.05$) between the per-

Table 5. Correlations among nations' percent girls on IMO team, GGI, team rank, and population (for nations with median team rank among top 30)

	% girls on IMO team 1989–2008	Median team rank 2000–2008	% of world population
Median team rank	−0.075	—	—
% of world population	−0.181	−0.333†	—
Gender gap index 2007	0.441*	0.167	−0.323†

†, $P < 0.10$; *, $P < 0.05$.

centage of girls on a country's IMO teams during the past 2 decades (Table 4) and its 2007 GGI (Table 5 and Fig. 3). These findings stand in distinct contrast to Machin and Pekkarinen's (19) claim that no correlation exists between GGI and VR for mathematics.

Noteworthy in this context is the fact that the U.S. ranked only 31st best, between Estonia and Kazakhstan, among the 128 countries included in the *2007 Global Gender Gap Report* (30). Countries such as the U.K. and Iceland, where the ratio of girls-to-boys scoring above the 99th percentile in the 2003 PISA was close to 1.0 or favored girls, had a GGI rank of 11 and 4, respectively. Likewise, Denmark and the Netherlands, where the VR was essentially 1.00, had GGI ranks of 8 and 12, respectively.

Similarly, Penner's cross-nation analysis of the 1995 TIMSS data (20) showed that the proportion of girls scoring above the 95th percentile positively and significantly correlated with several measures of female equality and status, including equity in educational opportunities and representation in the labor force and political offices. Numerous findings in Penner's recent study contradict the Greater Male Variability Hypothesis for mathematics. Our finding that the percentage of girls on a country's IMO team significantly correlates with its GGI, but not with its median IMO team rank or percentage of world population (Table 5), is also inconsistent with the Greater Male Variability Hypothesis. If this hypothesis were valid, these latter factors should inversely and significantly correlate with percentage of girls because (*i*) the best 6-member IMO teams consist of multiple students gifted at the 1-in-a-million level where females would be rare, and (*ii*) countries with larger populations would be more likely to have several such 5-SD-above-the-mean students; the GGI would be largely irrelevant.

Thus, we conclude that gender inequality, not greater male variability, is the primary reason fewer females than males are identified as excelling in mathematics at the high and highest levels in most countries. Of course, gender inequity is complex and multifaceted. It can encompass dynamics in school classrooms leading teachers to provide more attention to boys; guidance counselors, biased by stereotypes, advising females against taking engineering courses; mathematically gifted girls not being identified and nurtured; scarcity of women role models in math-intensive careers leading girls to believe they do not belong in them; unconscious bias against females in hiring decisions; and hostile work environments leading qualified women to drop out in favor of friendlier climes. The data reviewed here did not determine which of these and other gender-related factors are most influential; all likely contribute to some degree.

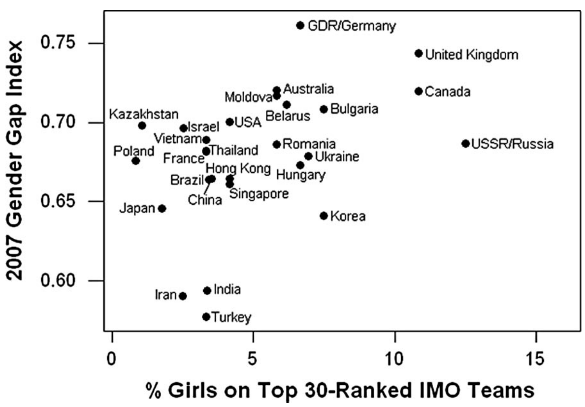

Fig. 3. Presence of females on top 30-ranked IMO teams strongly correlates ($r = 0.44$, $P < 0.05$) with measures of gender equity within countries. The IMO data for percentage of girls on countries' teams from 1989 to 2008 were taken from Table 4. The GGIs were taken from ref. 30.

Conclusions and Future Directions

This review was organized around 3 questions: Do gender differences in mathematics performance exist in the general population? Do gender differences exist among the highly mathematically talented? Do females exist who possess profound mathematical talent? The answer to the first question is that U.S. girls now perform as well as boys on standardized mathematics tests at all grade levels. Among the mathematically gifted, there may be as many as 2- to 4-fold more boys than girls depending on precisely where the cutoff is set. However, this gender gap, too, has been closing over time at all levels, including even in the IMO. Thus, there is every reason to believe that it will continue to narrow in the future. Moreover, the gender ratio favoring boys above the 99th percentile is not ubiquitous and correlates well with measures of a country's gender equity, strongly indicating that the gap is due, in large part, to sociocultural and other environmental factors, not biology or gender per se.

One serious policy concern that arose from the Hyde et al. study (13) is that the tests developed by states in the U.S. to comply with the mandates of NCLB include almost no questions requiring complex problem solving. NCLB puts pressure on teachers to try to get all their students to pass, thus leading them increasingly to teach to the test (32). With complex problem solving not covered, mathematics teachers will be tempted to neglect teaching it in favor of teaching computation and other lower-level mathematics skills. Yet problem solving and high-level mathematical reasoning are essential skills for success in life and STEM careers. This neglect of problem-solving skills could place U.S. students at a disadvantage compared with their peers in countries where teaching and tests emphasize more challenging content (33). Therefore, it is crucial to address this issue.

Importantly, the U.S. also needs to do a better job of identifying and nurturing its mathematically talented youth, regardless of their gender, race, or national origin. Doing so is vital to the future of the U.S. economy as documented in Thomas Friedman's *The World Is Flat: A Brief History of the Twenty-First Century* (34). *Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering* (35), *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future* (36), *Foundations For Success: The Final Re-*

port of the National Mathematics Advisory Panel (37), and *Identifying and Cultivating Extraordinary Mathematical Talent* (38) outline numerous steps the U.S. can and should take to ensure we have the well-educated labor force needed to fill the STEM jobs of the future.

ACKNOWLEDGMENTS. We thank Jonathan Kane and Jennifer Petersen for help with making figures and performing statistical analyses and Marcia Linn and Cathy Kessel for comments on the manuscript. This work was supported by National Science Foundation Grant REC 06354444 (to J.S.H.) and funds from The Wisconsin Alumni Research Foundation (to J.E.M.).

1. Shields S (1975) Functionalism, Darwinism, and the psychology of women. *Amer Psychol* 30:739–754.
2. Romanes G J (1887). Mental differences between men and women. *Nineteenth Century* 21:654–672.
3. Maccoby E (1966) *The Development of Sex Differences* (Stanford Univ Press, Stanford, CA).
4. Maccoby E, Jacklin C (1974). *The Psychology of Sex Differences* (Stanford Univ Press, Stanford, CA).
5. Spelke E (2005) Sex differences in intrinsic aptitude for mathematics and science? A critical review. *Amer Psychol* 60:950–958.
6. Lipsey MW, Wilson DB (2001) *Practical Meta-Analysis* (Sage, Thousand Oaks, CA).
7. Cohen J (1988) *Statistical Power Analysis for the Behavioral Sciences* (Lawrence Erlbaum, Hillsdale, NJ).
8. Hyde JS, Fennema E, Lamon S (1990). Gender Differences in Mathematics Performance: A meta-analysis. *Psychol Bull* 107:139–155.
9. Hedges LV, Nowell A (1995) Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science* 269:41–45.
10. Meece JL, Parsons JE, Kaczala CM, Goff SB (1982) Sex differences in math achievement: Toward a model of academic choice. *Psychol Bull* 91:324–348.
11. Eccles JS (1994) Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *Psychol Women Quart* 18:585–610.
12. National Science Foundation (2006) *Science and Engineering Indicators 2006.* Available at www.nsf.gov/statistics/seind06.
13. Hyde JS, Lindberg SM, Linn MC, Ellis AB, Williams CC (2008) Gender similarities characterize math performance. *Science* 321:494–495.
14. National Assessment of Educational Progress (NAEP) (2008) *The Nation's Report Card*. Available at http://nces.ed.gov/nationsreportcard/itmrls/startsearch.asp.
15. Weinberger CJ (2005) *Is Science and Engineering Workforce Drawn from the Far Upper Tail of the Math Ability Distribution?* Available at http://econ.ucsb.edu/~weinberg/uppertail.pdf.
16. Hedges LV, Friedman L (1993) Sex differences in variability in intellectual abilities: A reanalysis of Feingold's results. *Rev Educ Res* 63:94–105.
17. National Science Foundation (2008) *Women, Minorities, and Persons with Disabilities in Science and Engineering*. Available at www.nsf.gov/statistics/wmpd/pdf/tabf-2.pdf.
18. Guiso L, Monte F, Sapienza P, Zingales L (2008) Culture, gender, and math. *Science* 320:1164–1165.
19. Machin S, Pekkarinen T (2008) Global sex differences in test score variability. *Science* 322:1331–1332.
20. Penner AM (2008) Gender differences in extreme mathematical achievement: An international perspective on biological and social factors. *Am J Sociology* 114:S138–S170.
21. Benbow CP, Stanley JC (1983) Sex differences in mathematical reasoning ability: More facts. *Science* 222:1029–1031.
22. Brody L, Mills C (2005) Talent search research: What have we learned? *High Ability Stud* 16:97–111.
23. Monastersky R (2005) Women and Science: The debate goes on. *Chronicle Higher Educat* 51(26):A1.
24. Green J (2000) How Many Women Mathematicians Can You Name? Available at www.users.muohio.edu/porterbm/sumj/2000/Green00.pdf
25. Green J, LaDuke J (2009) Pioneering Women in American Mathematics: The Pre-1940s PhD's. In *History of Mathematics* (American Mathematical Society, Providence, RI), Vol. 34.
26. Phipps P, Maxwell JW, Rose C (2009) 2008 annual survey of the mathematical sciences (first report) preliminary report on the 2007–2008 new doctoral recipients. *Notices AMS* 56:257–267.
27. Nelson DJ (2007) *A National Analysis of Diversity in Science and Engineering Faculties at Research Universities.* Revised October 20, 2007. Available at http://chem.ou.edu/~djn/diversity/briefings/Diversity%20Report%20Final.pdf.
28. Andreescu T, Gallian JA, Kane JM, Mertz JE (2008) Cross-cultural analysis of students with exceptional talent in mathematical problem solving. *Notices AMS* 55:1248–1260.
29. Baker DP, Jones DP (1993) Creating gender equality: Cross-national gender stratification and mathematical performance. *Sociol Educ* 66:91–103.
30. Hausmann R, Tyson LD, Zahidi S (2006,2007) *The Global Gender Gap Report* (World Economic Forum, Geneva, Switzerland). Available at www.weforum.org/pdf/gendergap/report2007.pdf.
31. National Center for Education Statistics (2007) *Digest of Education Statistics.* Available at http://nces.ed.gov/programs/digest/d07.
32. Au W (2007) High-stakes testing and curricular control: A qualitative metasynthesis. *Educ Res* 36:258–267.
33. Roth K, Garnier H (2006) What science teaching looks like: An international perspective. *Educ Leadership* 64:16–23.
34. Friedman TL (2005) *The World Is Flat: A Brief History of the Twenty-First Century* (Farrar, Straus, and Giroux, New York, NY).
35. Committee on Maximizing the Potential of Women in Academic Science and Engineering and Committee on Science, Engineering, and Public Policy (2006) *Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering. National Academy of Sciences, National Academy of Engineering, and Institute of Medicine of the National Academies* (The National Academies Press, Washington, DC). http://books.nap.edu/catalog.
36. Committee on Science, Engineering, and Public Policy (2006) *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future. National Academy of Sciences USA* (The National Academies Press, Washington, DC). Available at http://books.nap.edu/openbook/0309100399/html/index.html.
37. National Mathematics Advisory Panel (2008) *Foundations for Success: The Final Report of the National Mathematics Advisory Panel* (U.S. Department of Education, Washington, DC). Available at www.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf.
38. Andreescu T, Gallian JA, Mertz JE (2008) *Teachers College Record, Nov. 26, 2008.* Available at www.tcrecord.org.