Explaining the Gender Gap in Math Test Scores: The Role of Competition

Muriel Niederle and Lise Vesterlund

December 16, 2009

Muriel Niederle is Associate Professor of Economics, Stanford University, Stanford, California. She is also a Research Associate, National Bureau of Economic Research, Cambridge, Massachusetts. Lise Vesterlund is the Andrew W. Mellon Professor of Economics, University of Pittsburgh, Pittsburgh, Pennsylvania. Their e-mail addresses are <niederle@stanford.edu> and <vester@pitt.edu>, respectively.

Over the past 60 years there have been substantial improvements in the college preparation of female students, and the college gender gap has changed dramatically. Goldin, Katz and Kuziemko (2006) show that female high school students now outperform male students in most subjects, and in particular on verbal test scores. The ratio of male to female college graduates has not only decreased, but reversed itself, and the majority of college graduates is now female.

The gender gap in mathematics has also changed. The number of math and science courses taken by female high school students has increased and now the mean and standard deviation in performance on math test scores are only slightly larger for males than for females. Despite minor differences in mean performance, Hedges and Nowell (1995) show that many more boys than girls perform at the right tail of the distribution. This gender gap has been documented for a series of math tests including the AP calculus test, the mathematics SAT, and the quantitative portion of the GRE. Over the past 20 years the fraction of males to females who score in the top five percent in high school math has remained constant at two to one (Xie and Shauman, 2003). Examining students who scored 800 on the math SAT in 2007, Ellison and Swanson (this issue) also find a two to one male-female ratio. Furthermore, they find that the gender gap widens dramatically when examining the right tail of the performance distributions for students who participate in the American Mathematics Competitions (AMC).

Substantial research has sought to understand why more boys than girls excel in math. However, given the many dimensions in which girls outperform boys, it may seem misplaced to focus on the dimension in which girls are falling short. Why not examine the gender gap in verbal test scores where females outperform males? One reason is that in contrast to, say, verbal test scores math test scores serve as a good predictor of future income. Although the magnitude of the effect of math performance on future income varies by study, the significant and positive effect is consistently documented, see e.g., Paglin and Rufolo (1990), Murnane, Willet and Levy (1995), Grogger and Eide (1995), Weinberger (1999, 2001), Murnane, Willett, Duhaldeborde and Tyler (2000) and Altonjii and Blank (1999) for a discussion.

So why do girls and boys differ in the likelihood that they excel in math? One argument is that boys have and develop superior spatial skills, and this gives them an advantage in math. Such differences could have an evolutionary foundation, as male tasks such as hunting may have required greater spatial orientation than typical female tasks (for a discussion see Gaulin and Hoffman, 1988). In addition, or alternatively, it could be because boys tend to engage in play that is more movement oriented and therefore grow up in more spatially complex environments (Berenbaum et al., 2008).

The objective of this paper is not to discuss whether the mathematical skills of males and females differ, be it a result of nurture or nature. Rather we argue that the reported test scores need not reflect the magnitude of the gender differences in math skills. We will present results that suggest that the abundant and disturbing evidence of a large gender gap in mathematics performance at high percentiles in part may be explained by the differential manner in which men and women respond to competitive test taking environments.

We provide evidence of a significant and substantial gender difference in the extent to which skills are reflected in a competitive performance. The effects in mixed-sex settings range from women failing to perform well in competitions (Gneezy, Niederle and Rustichini, 2003) to women shying away from environments in which they have to compete (Niederle and Vesterlund, 2007). We find that the response to competition differs for men and women, and in the examined environment gender difference in competitive performance does not reflect the difference in non-competitive performance.

We use the insights from these studies to argue that the competitive pressures associated with test taking may result in performances that do not reflect those of less competitive settings. Of particular concern is that the distortion is likely to vary by gender and that it may cause gender differences in performance to be particularly large in mathematics and for the right tail of the performance distribution. Thus the gender gap in math test scores may exaggerate the math advantage of males over females. Due to the way tests are administered and rewards are allocated in academic competition, there is reason to suspect that females are failing to realize their full potential or to have that potential recognized by society.

Gender Differences in Competitive Performance and Selection

Performance in Competitive Environments

Clear evidence that incentive schemes may generate gender differences in performance has been shown by Gneezy, Niederle and Rustichini (2003). In an experiment conducted at the Technion in Israel, individuals are presented with an incentive scheme and asked to solve mazes on the Internet for 15 minutes. Four different incentive schemes were examined. Thirty women and thirty men perform under each incentive scheme, with no one performing under more than one incentive scheme. Though gender was not explicitly mentioned, participants could see one another and determine the gender composition of the group.

In a non-competitive environment, three men and three women receive an individual piece-rate payment of $0.50 for every maze he or she solves. In this environment, the gender gap in performance is small, with men solving an average of 11.2 mazes and women solving 9.7 mazes. The emphasis is not on determining whether this gender gap in performance reflects differences in ability, experience or performance costs, but rather on determining how the gender gap responds to an increase in competition. That is, will the performance gap seen in a competitive environment reflect the gap seen in this non-competitive piece-rate environment? To examine performance under competitive pressure Gneezy et al. (2003) ask a different set of participants to compete in groups of three men and three women under a tournament incentive scheme. The participant with the highest performance in each group receives a payment of $3 per maze, while the other members of the group receive no payment. Compared to the piece-rate the mixed-sex tournament significantly increases the average performance of men while that of women is unchanged. This creates a significant gender gap in performance of 4.2 mazes, which substantially exceeds the average performance difference of 1.5 in the non-competitive environment. Thus the gender gap in performance under competition is three times greater than that seen under the piece-rate payment. Results are summarized in Figure 1, first showing the gender-gap in performance in the piece rate and then in the mixed-sex tournament.

<Insert Figure 1 about here>

Differences in performance between the piece rate and the tournament can stem from the introduction of competition, but also from the fact that the tournament compensation is more uncertain. To determine whether the differential response to competition is driven by gender differences in risk aversion, a random-pay scheme was implemented where one member of each group (three men and three women) was selected randomly after the performance to receive a payment similar to the tournament payment of $3 for every maze solved. The payment for everyone else was zero. If gender differences in risk aversion played a substantial role in explaining the behavior in mixed-sex tournaments then we would expect the random-pay treatment to generate a large gender difference in performance as well.[1] In contrast, Figure 1 shows that the average performance gap under random pay is similar to the one in the piece rate.

A final treatment examines performance in single-sex tournaments, with six men or six women in each group. In this case both men and women improve their performance compared to non-competitive incentive schemes. The resulting gender gap in mean performance is 1.7 in the

[1] Eckel and Grossman (2008) and Croson and Gneezy (2009) summarize the experimental literature in economics and conclude that women exhibit greater risk aversion. Byrnes, Miller and Shafer (1999) present a meta-analysis of 150 psychology studies and demonstrate that while women in some situations are significantly more averse to risk, many studies find no gender difference.

single-sex tournament, which is similar to the gaps of 1.5 in the piece-rate and the random-pay treatment, but much smaller than the 4.2 gap in the mixed-sex tournaments. The gap in the mixed-sex tournament is significantly higher than in the three other treatments. Hence, it is not the case that these women at the Technion generally are unwilling or unable to perform well in competitions, but rather that they do not compete well in competitions against men.[2]

How does competition influence the gender composition of the top performers? Due to the number of subjects the top two quintiles are examined, i.e., the best forty percent of performers. In both of the non-competitive treatments and in the single-sex tournament women account for 40 percent of those in the top two quintiles. Thus if the tournaments were run in single-sex groups one may falsely conclude that men and women have similar responses to competition. However, running mixed-sex tournaments significantly decreases the fraction of women with a performance in the top two performance quintiles from 40 to 24 percent. Thus in mixed-sex competitions we see a decrease in the relative performance of women and in the fraction of women in the top two performance quintiles.

Entering Competitions

If women are uncomfortable performing in a competitive setting, then they may be less likely to enter competitive settings. Niederle and Vesterlund (2007) examine whether men and women differ in their willingness to enter a mixed-sex competition. Forty men and forty women from the subject pool at the Pittsburgh Experimental Economic Lab participated in the experiment. Participants were asked to add up sets of five two-digit numbers for five minutes under different compensation schemes. At the end of the experiment one performance was selected to count for payment. Participants were paid for their performance, which was measured in each task by the number of correctly solved problems. No participant was restricted in the number of problems that could be solved. Participants were not informed of the performance by anyone else until the end of the study, and were told of each compensation scheme only immediately before performing the task.

Participants first perform the task under a non-competitive piece rate where they received 50 cents per correctly solved problem. Subsequently they perform in tournaments of two men and two women. While gender was never mentioned during the experiment, individuals could see

[2] Gneezy and Rustichini (2004) document results in 40 meter running competitions among 10 year-olds. Children first run 40 meters separately, and then compete against another child with a similar performance. They find no initial gender difference in speed. However, in general boys win the competition against girls independent of the girl's initial performance. In same-sex competitions the likelihood of winning the competition is almost the same for the faster child as it is for the slower child.

their competitors and determine the gender composition of the group. Only the person with the largest number of correctly solved problems was paid and received $2 per correct problem. The other members of the group received no payment. Under the piece rate men and women solve an average of 10.7 and 10.2 problems, respectively, and under the tournament they solve 12.1 and 11.8, respectively. Neither case demonstrates a significant gender difference in performance. Thus, for this very short task of simple math problems men and women did not differ in their ability to compete in mixed-sex groups. In fact, for this specific short task, changes in incentives do not appear to have a large effect on performance. Later examinations suggest that the increase in performance from the piece rate to the tournament is driven largely by experience.

Having performed both under the piece rate and the tournament compensation scheme, participants were asked which of the two they would prefer for their performance on a subsequent five-minute addition task. To secure that the individual's choice only depends on the participant's beliefs on relative performance, we designed the choice as an individual decision. Specifically, a participant who selected the tournament would win if his or her new performance exceeded the performance of the three other group members from the previous competition.

Given the lack of a gender gap in performance, maximization of earnings predicts no gender difference in choice of compensation scheme. In contrast to the prediction we observe a substantial gender gap in tournament entry. Seventy-three percent of the men and 35 percent of the women enter the tournament.[3]

Figure 2A shows the proportion of men and women who enter the subsequent tournament for each initial tournament performance quartile. Neither the tournament-entry decisions of men nor those of women are very sensitive to the individual's performance, and independent of the performance quartile, men are much more likely to enter the tournament. On average men in the worst performance quartile enter the tournament more than women in the best performance quartile.

<Insert Figure 2 about here>

To study the effect of beliefs about relative performance, participants were asked to rank their performance in the initial tournament. Any correct guess was rewarded by $1. Accounting for ties, at most 30 percent of men and women should guess that they are the best in their group

[3] A gender gap in willingness to compete has also been documented by Niederle, Segal and Vesterlund (2008), Dargnies (2009), Cason, Masters and Sheremeta (2009), Gneezy and Rustichini (2005), Gupta, Poulsen and Villeval (2005), Prize (2008a) and Wozniak (2009). Gneezy, Leonard and List (2009) replicate the finding in a patriarchal African society, but not in a matrilineal Indian one. Prize (2008b) examines men and women who are equally confident and find that there is no gender difference in competitive entry.

of four. We find that 75 percent of men compared to 43 percent of women guessed that they were the best. While both men and women are overconfident, men are more overconfident than women. Figure 2B shows that while beliefs predict tournament entry for both men and women a substantial gender gap in entry remains. Among those who reported that they thought they were best in their group of four, 80 percent of men enter the tournament compared to only 50 percent of women. This 30 percentage point gender gap in tournament entry remains among those who thought they were second out of four. With 84 percent of participants reporting that they were ranked first or second, it follows that there is a substantial gender gap in competitive entry even conditional on beliefs. Regressions confirm this result when controlling for both performance and beliefs.

Other possible reasons for the different compensation choices of men and women may be that they differ in their attitudes toward risk and feedback on relative performance. The compensation scheme associated with the tournament is more risky and results in the participant receiving feedback on relative performance. In our study we find little evidence that these factors play a large role in explaining gender differences in tournament entry.[4] Controlling for the effects of beliefs, risk and feedback aversion, there remains a substantial and significant gender difference in tournament entry. We attribute this remaining difference to men and women differing in their attitude towards placing themselves in environments where they have to compete against others.

Our results show that women shy away from competition while men embrace it, and this difference is explained by gender differences in confidence and in attitudes toward competition. A consequence is that from a payoff-maximizing perspective too few high-performing women and too many low-performing men enter the tournament. Perhaps most important is that the fraction of women who win the competitions drops dramatically. When women had no option but to compete in randomly generated groups they are predicted to win 48 percent of competitions, however if competitions were run solely among those who opt to compete we instead predict that 29 percent of competitions would be won by women. Thus selection alone causes very few women to win competitions

Taking these studies together, the evidence suggests that in mixed-sex environments where there appear to be no or small gender differences in non-competitive performance, men nonetheless outperform women in competitions, and more frequently select a competitive

[4] The evidence on the extent to which gender differences in tournament entry is explained by gender differences in risk attitudes is mixed (for example, Cason, Masters and Sheremeta, 2009; Gupta, Poulsen and Villeval, 2005; Dohmen and Falk, 2006).

compensation. We can draw a strong parallel between the two research findings by interpreting the lower performance of women in the mixed-sex tournaments in Gneezy, Niederle and Rustichini (2003) as women choosing not to compete. The high female performance in the single-sex tournament shows that it is possible for women to perform well in competitions. However, the results of both studies suggest that women may not perform to their maximal ability in mixed-sex competitions.

The Effect of Competition on Math Test Scores

While test scores traditionally were thought to measure an individual's cognitive ability, researchers have come to recognize that test scores are influenced by cognitive as well as non-cognitive abilities (for example, Cunha and Heckman, 2007; and Segal, 2008). In particular, non-cognitive factors such as motivation, drive, and obedience may not only affect an individual's investments in cognitive skills, but also the individual's test score performance. In a nice demonstration of the effect of incentives on performance, Gneezy and Rustichini (2000) have participants solve a 20-minute IQ test under varying incentive schemes. They show that performance is lower when individuals are given a low piece rate per correct answer, rather than a high piece rate or even zero payment. Thus, students who have similar skills may receive different test scores if the incentives associated with a high performance differ or are perceived to differ. This suggests that test scores may reflect much more than cognitive skills.

A non-cognitive skill which may influence test scores is an individual's response to competitive pressure. The studies described above show that men and women differ in their response to competition when performing in mixed-sex environments. Thus, a very competitive test may result in gender differences in test scores that need not reflect the magnitude or the direction of gender differences in performance seen in less competitive environments.

Örs, Palomino and Peyrache (2008) elegantly show the relevance of this point in practice. They examine the performance of women and men in an entry exam to a very selective French business school (HEC) to determine whether the observed gender differences in test scores reflect differential responses to competitive environments rather than differences in skills. The entry exam is very competitive: only about 13 percent of candidates are accepted. Comparing scores from this exam reveals that the performance distribution for males has a higher mean and fatter tails than that of females. This gender gap in performance is then compared both to the outcome of the national high school exam, and for admitted students to their performance in the first year. While both of these performances are measured in stressful environments, they are much less competitive than the entry exam. The performance of women is found to dominate that of men, both on the high school exam and during the first year at HEC. Of particular interest is that females from the same cohort of candidates performed significantly better than males on the national high school graduation exam two years prior to sitting for the admission exam. Furthermore, among those admitted to the program they find that within the first year of the M.Sc. program females outperform males. Caution should however be used when comparing these results to those on the entry exam, not only is this a truncated sample of the original distribution, it is also one from which certain students may have exited. The authors also control for explanations pertaining to risk aversion and specific test-taking strategies. They find that for each student the variance of grades across different subjects is not higher for male than female students. This excludes a difference in strategies where a student studies a few topics intensively rather than studying all topics on a subject. Furthermore, they show that the same differences arise when focusing separately on the math and non-math part of the exam. They conclude that the differences in the gender gap between the entry exam and the high school exam as well as the first-year performance result from men and women differing in their response to competition.

Although no comparable study has been conducted in the United States, Örs et al (2008) note that their results are consistent with the observation that female grade point averages in both high school and college exceed those of males when controlling for their SAT scores (for example, Rothstein, 2004).

These findings suggest that caution is needed when using test scores to infer gender differences in skills. However, it is not clear why this should be more of an issue when looking at math rather than say verbal test scores or why a bias in math may be exacerbated at the right tail of the distribution? We will argue that the gender differences that were found to play an important role in the Niederle and Vesterlund (2007) study, confidence and attitudes towards competition, are likely to influence performance on competitive math tests and that these differences may play a substantial role at the right tail of the distribution.

Confidence, Stereotypes, and Math Tests

We begin by discussing why gender differences in confidence may be particularly large in mathematics. Girls and boys with the same math test scores have very different assessments of their relative ability (for example, Eccles, 1998). Conditional on math performance, boys are more overconfident than girls, and this gender gap is greatest among gifted children (Preckel, Goetz, Pekrun and Kleine, 2008). The strong gender stereotype that boys are better at math may help to explain this gender gap in confidence. This stereotype is further re-enforced by the fact that the fraction of male teachers in math-intensive courses is higher than for other classes.[5] Another source through which stereotypes may affect beliefs is shown by Jacobs (1991), who found that mothers who endorsed a male-math stereotype underestimated their daughters' ability in math. These perceptions were shown to be particularly important for a child's confidence because a child's self-evaluation of academic competency appears to be more strongly related to their parents' appraisals of their academic ability than to their actual academic performance.[6]

The findings by Pope and Sydnor (this issue) are very much in line with stereotypes influencing test performance at the tail. Looking at U.S. data they find large variation in the gender ratios of 8th graders scoring in the top 75th and 95th percentiles of the National Assessment of Educational Progress (NAEP). The test is taken by a sample of children in public schools. Consistent with beliefs influencing behavior, they show that in regions where men and women are viewed as more equal there are smaller gender disparities in stereotypically male-dominated tests of math and science and in stereotypically female-dominated tests of reading.

The relationship between perception of women and the math performance gap has also been documented across OECD countries. Guiso, Monte, Sapienza and Zingales (2008) use the 2003 Programme for International Student Assessment (PISA) evaluating 15-year-old students from 40 countries in identical tests in mathematics and reading. The tests were designed by the OECD to be free of cultural biases. They use several measures for the gender equality of a country, including the World Economic Forum's Gender Gap Index (Hausman, Tyson and Zahidi, 2006). Examples of European countries with high GGI scores are countries like Sweden, Finland and Norway, while low GGI countries are France, Greece and Italy. In countries that score highly on gender equality Guiso et al (2008) find a smaller gender gap in mean math performance as well as in the tail of the distribution. In contrast to Pope and Sydnor (this issue) they find a positive correlation between math and reading with women performing well on both tasks in societies with greater gender equality.

---

[5] Dee (2007) and Carrell, Page and West (2009) study the effect of a teacher's gender on performance. Having a female math or science teacher improves the math and science performances by females, and the effect is particularly large for the gifted female students. Using the 1999-2000 Schools and Staffing Survey (SASS) Dee (2007) estimates that in 12th grade 44 and 52 percent of science and math teachers are female, compared to 71 percent in reading. See Bettinger and Long (2005) for evidence on college instruction.

[6] Stereotypes may not only influence a child's confidence directly and the manner in which the child responds to competition, it may also influence the likelihood by which the child "chokes" in any performance setting. *Stereotype threat theory* (Steele, 1997) argues that a strong stereotype may harm the stereotyped individual's performance on a task because they fear confirming it. Spencer et al. (1999) show that the effect of stereotype threat may be removed if in describing a test it is stated that the "math test had revealed no gender difference in the past."

Looking at the very highest performing women in mathematics, Hyde and Mertz (2009) examine the proportion of women among delegates at the International Mathematical Olympiad (IMO) in the last two decades for countries that achieved a median rank among the top 30 in recent years. The proportion of females in a countries' IMO team is not correlated with median team rank. However, they find a positive correlation between the percentage of girls in a country's IMO team during the past two decades and its 2007 Gender Gap Index. Ellison and Swanson (this issue) do not replicate this finding and argue that it may be because they examine a larger set of countries. They note that when examining the very high achieving students the gender gap is very large and particularly troubling is that top performing girls in this set are concentrated in a few elite schools compared to the top performing boys.

The strong stereotype of male superior math performance may influence the confidence of females and affect their performance on competitive math tests. This effect is likely to be exacerbated for those at the tail of the distribution for whom the gender gap in confidence may be large.

Attitudes Towards Competition and Math Tests

Why might gender differences in competitive attitudes be more of an issue on math tests? One reason may be that math answers are either right or wrong, thus in contrast to verbal test scores, math test scores may better predict actual rank as well as future relative performances. Another reason is that more boys select math-intensive majors, which in turn increases the fraction of relevant male competitors on math tests relative to that on say verbal tests. As shown by Gneezy et al. (2003), a woman's competitive performance is sensitive to the gender of her competitors. While women at the Technion improved their performance when competing in all-female groups, this was not the case in mixed-sex groups.

To parse the effects of the gender composition of competitors, Niederle, Segal and Vesterlund (2009) extend the original Niederle and Vesterlund (2007) study. The initial finding that gender differences in confidence and attitudes toward competition help explain tournament entry led us to examine the compensation choices of men and women in an "affirmative-action" tournament, where for every two winners we require that at least one winner must be a woman. Such a requirement not only increases the probability that women will win the tournament, it also makes the competition more gender-specific. In the affirmative-action tournament, a woman will win the competition if she is either the best performing woman or has one of the two highest performances in the group, a man on the other hand will have to both be the best-performing man and have one of the two highest performances in the group. Increasing the number of same-sex competitors may affect the decision to enter a tournament because both the gender gap in beliefs as well as in attitudes to competition could be smaller in more gender-specific competitions. If women are more comfortable competing against women, this may influence their compensation choices.

The experiment was conducted at the Harvard Business School, using students from the Computer Lab for Experimental Research (CLER) lab subject pool. Participants in the experiment compete in groups of three men and three women. They are presented with two different compensation choices. In the standard tournament choice they choose between a 50-cent piece rate and a tournament where the two participants with the largest number of correctly solved problems each will be paid $1.50 per correctly solved problem and the remaining four members will receive no payment. In the second choice participants instead choose between a 50-cent piece rate and a $1.50 affirmative-action tournament. The two winners of the affirmative-action tournament are the highest performing woman and the highest performer of the remaining five members of the group.

Our study shows that when women are guaranteed equal representation among winners, more women and fewer men enter competitions and the change exceeds that predicted by the changes in the probability of winning that result from the introduction of affirmative action. The response causes the fraction of entrants who are women to increase from 29 to 64 percent. The excessive response is explained to a large extent by changes in beliefs on the chances of winning the competition and attitudes toward competition. Specifically, men are less overconfident and women less reluctant to compete in groups where their own gender is better represented.

The sensitivity to gender composition is also shown by Huguet and Regner (2007). When girls are led to believe that a task measures math ability then they are found to underperform in mixed-sex groups, but not in all-female groups. A natural question may be why women are more apprehensive toward competitions against males. One explanation may be that it is more threatening to compete against individuals who are overconfident and very eager to compete and win.

The reported studies suggest that a woman’s performance and willingness to compete is sensitive to the gender of those she is competing with. If a large fraction of competitors on math tests are male, then gender differences in attitudes toward competition may play a particularly large role, and this effect may be exacerbated at the more male dominated upper tail.

CONCLUSIONS

A series of studies have shown that males and females differ in their response to competition. We have argued that such gender differences may cause test scores to magnify and potentially distort underlying gender differences in skills. In light of the role played by beliefs on relative performance and women's sensitivity to competition against men, these factors may be particularly important when assessing math skills.

The reported studies suggest that competitive pressure may cause gender differences in test scores that exaggerate the underlying gender differences in math skills. Needless to say this distortion is not a concern if an individual's test score is not simply meant to reflect math skills, but rather math skills under competitive pressure. Certainly math test scores may be very good predictors of winners of the American Mathematics Competition. However there are many circumstances where math test scores are solely used to assess math skills. In those situations we may need to be cautious of the bias that the competitive environment imposes on women.

We have focused on explaining how a differential response to competition may distort gender differences in test scores. However, sensitivity to competitive pressure is also likely to influence the investment in and selection into male dominated or math intensive fields where there are strong stereotypes on female inabilities. If educational investments vary by gender, and these influence a student's preparedness when taking the test, then this may further explain the differences in math test scores.

At the high school level there is little evidence that girls on average invest less in math than boys. Goldin et al. (2007) show that girls and boys take advanced math classes at similar rates, and Guiso et al. (2008) find that if anything girls spend more time on math homework than boys. While these studies demonstrate that on average there are no gender differences in math skill investments, it would be of interest to determine whether the same holds at the upper tail of the distribution.

At the college level there are substantial gender differences in math related investment. It is important however to note that these investments need not reflect differences in skills. In an experiment using Stanford undergraduates Niederle and Yestrumskas (2008) show that females may be less likely to choose a difficult task. They first have women and men solve an easy task. When asked to choose the difficulty for a subsequent performance task, men select a challenging task over an easy task 50 percent more often than women, even when controlling for initial performance and beliefs about one's performance. This result is consistent with those of Lefevre et al. (1992) and Weinberger (2005) who find that among equally gifted students, males are many times more likely to select college majors that are considered to be high in math content. Furthermore the drop-out rate for these majors is much greater for women.

Many factors may explain why fewer women end up completing math intensive college course work. Partial explanations may be found in examining the explanations women give for dropping out of these courses. A report entitled "Women's Experiences in College Engineering," funded by the National Science Foundation and the Sloan Foundation, writes that the exit of many young women is not driven by ability, but rather that this decision is influenced by women negatively interpreting their grades and having low self-confidence (Goodman, Cunningham, and Lachapelle, 2002). Furthermore these women mention that negative aspects of their schools' climate, such as competition, lack of support and discouraging faculty and peers, cause them to reevaluate their field of study. In an earlier study of engineering student performance and retention, Felder et al. (1995) find similar effects.

A crucial question is whether it is possible to alter how women perceive and experience math intensive studies. Advocates for single-sex education have long argued that the gender composition in the class room can influence a girl's investment in both math and science. Indeed the evidence presented in a recent study by Fryer and Levitt (2009) suggests that single-sex education may improve the confidence of girls and cause them to hold less stereotypical views of gender roles. In a cross-country analysis Fryer and Levitt find that there is no gender gap in math performance in Middle Eastern countries with same-sex schooling, and this causes them to speculate that perhaps the relationship between single-sex schooling and the absence of a gender gap in math performance may be causal. Unfortunately it is difficult to use cross-country data to determine the effect of single-sex education. The U.S. evidence of single-sex schooling is far from conclusive, while some find that girls from single-sex schools are more likely to subsequently enter sciences, others fail to find such an effect (for a discussion see Campbell and Sanders, 2002). As single-sex schools in the US are private self-selection may play a significant role and identification of an effect of single-sex schooling on math achievements is difficult.[7]

In conclusion, gender differences in competitive attitudes may cause mathematics test scores to give a biased representation of the underlying gender differences in math skills. Our results suggest that it may be important to examine whether changes in testing or evaluation may allow more females to realize their potential and better measure their current math interests and math skills.

[7] Booth and Nolen (2009) examine tournament-entry decision by boys and girls in mixed- and single-sex schools. They find that girls from selective single-sex schools are more likely to enter competitions against boys than girls from non-selective mixed-sex schools. Unfortunately it is not possible to determine whether this response results from the superior performance by students at the single-sex schools.

References

Altonji, Joseph G., and Rebecca Blank, "Race and Gender in the Labor Market," in Handbook of Labor Economics, Vol. 3c, O. Ashenfelter, and D. Card, eds. (Amsterdam, The Netherlands: Elsevier Science B.V., 1999), 3144–3259.

Berenbaum, S. A., Martin, C. L., Hanish, L. D., Briggs, P. T., & Fabes, R. A. (2008). Sex differences in children's play. In J. B. Becker, K. J. Berkley, N. Geary, E. Hampson, J. Herman, & E. Young (Eds.), Sex differences in the brain: From genes to behavior. New York: Oxford University Press.

Bettinger, Eric and Bridget Long (2005) "Female Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students," *American Economic Review Papers and Proceedings* 95(2): 152-157.

Booth, Alison L. and Patrick J. Nolen, 2009, "Choosing to Compete: How Different Are Girls and Boys?", IZA working paper No. 4027.

Byrnes, James P., David C. Miller, and William D. Schafer, "Gender Differences in Risk Taking: A Meta-Analysis," *Psychological Bulletin*, LXXV, 1999, 367-383.

Campbell, Patricia B and Sanders, Jo. "Challenging the System: Assumptions and Data Behind the Push for Single-Sex Schooling," A. a. H. Datnow, Lea, Gender in Policy and Practice: Perspectives on Single-Sex and Coeducational Schooling. Routledge Falmer, 2002.

Carrell Scott E, Marianne E. Page and James E. West, 2009. "Sex and Science: How Professor Gender Perpetuates the Gender Gap," NBER Working Paper.

Cason, T.N.; Masters, W.A. and Sheremeta, R.M. "Entry into Winner-Take-All and Proportional-Prize Contests: An Experimental Study," 2009.

Croson, R. and Gneezy, U. "Gender Differences in Preferences." *Journal of Economic Literature*, 2009, 47(2), pp. 1-27.

Cunha, Flavio and James Heckman, 2007, "The Technology of Skill Formation," *American Economic Review Papers & Proceedings*, 97(2): 31- 47, (2007).

Dargnies, Marie-Pierre. "Does Team Competition Eliminate the Gender Gap in Entry in Competitive Environments?," Working Paper. CES, 2009.

Dee, Thomas S. 2007. "Teachers and the Gender Gaps in Student Achievement." *Journal of Human Resources* 42(3): 528–554.

Dohmen, Thomas and Armin Falk, 2006, "Performance Pay and Multi-dimensional Sorting: Productivity, Preferences and Gender ," working paper IZA DP 2001.

Eccles, Jacquelynne S. 1998. "Perceived control and the development of academic motivation: Commentary." *Monographs of the Society for Research in Child Development*, 63(2/3), 221-231. http://www.psc.isr.umich.edu/pubs/abs/1707

Eckel, Catherine C. and Philip J. Grossman, 2008, "Sex and Risk: Experimental Evidence," forthcoming, *Handbook of Experimental Economics Results*, Amsterdam, Elsevier Science, North-Holland, 1061-73.

Ellison, Glenn and Ashley Swanson, "The Gender Gap in Secondary School Mathematics at High Achievement Levels: Evidence from the American Mathematics Competitions", JEP this issue.

Felder, Richard; Felder, Gary; Mauney, Merdith; Hamrin, Charles and Dietz, Jacquelin. "A Longitudinal Study of Engineering Student Performance and Retention III: Gender Differences in Student Performance and Attitudes." *Journal of Engineering Education*, 1995, *84*(2), 151-163.

Fryer, Roland, and Steven Levitt, "An Empirical Analysis of the Gender Gap in Mathematics," August, 2009, forthcoming in *American Economic Journal: Applied Economics*.

Gaulin, Steven, and Harol Hoffman, (1988), "Evolution and development of sex differences in spatial ability", in *Human Reproductive Behavior: A Darwinian Perspective*, edited by

Laura Betzig, Monique Borgerhoff Mulder and Paul Turke, Cambridge University Press 1988.

Gneezy, Uri; Leonard, Kenneth L. and List, John A. "Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society," *Econometrica*, Vol. 77, No. 5 September 2009, 1637–1664.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini, “Performance in Competitive Environments: Gender Differences,” *Quarterly Journal of Economics*, CXVIII, August 2003, 1049 – 1074.

Gneezy, Uri, and Aldo Rustichini (2000) “Pay Enough or Don't Pay At All.” *Quarterly Journal of Economics* August, 791-810.

Gneezy, Uri, and Aldo Rustichini “Gender and competition at a young age,” *American Economic Review Papers and Proceedings*, May 2004, 377-381.

Gneezy, Uri and Aldo Rustichini, “Executive versus Teachers: Gender, Competition and Self-Selection”, November 2005, working paper.

Goldin, Claudia, Lawrence F. Katz and Ilyana Kuziemko, “The Homecoming of Amercian College Women: The Reversal of the College Gender Gap”, Journal of Economic Perspectives, vol. 20, No 4, Fall 2006, 133-156.

Goodman, I.F.; Cunningham, C.M. and Lachapelle, C. "The Women's Experience in College Engineering." *Goodman Research Group, Inc*, 2002. At <http://www.grginc.com/WECE_FINAL_REPORT.pdf>.

Grogger, Jeff, and Eric Eide. 1995. "Changes in College Skills and the Rise in the College Wage Premium." *Journal of Human Resources* 30(2):280-310.

Guiso, Luigi, Ferdinando Monte, Paolo Sapienza and Luigi Zingales, 2008, “Culture, Gender and Math, *Science*, May 30, Vol 320, 1164-1165.

Gupta, Nabanita Datta, Anders Poulsen and Marie-Claire Villeval, “Male and Female Competitive Behavior – Experimental Evidence”, November 2005, GATE working paper, Ecully, France.

Hausman, Ricardo Hausmann, Laura D. Tyson and Saadia Zahidi, 2006, *The Global Gender Gap Report*, World Economic Forum, Geneva, Switzerland.

Hedges Larry V. and Amy Nowell, “Sex differences in mental test scores, variability, and numbers of high-scoring individuals”, Science 1995: Vol. 269. no. 5220, pp. 41 – 45.

Huguet, P., Régner, I. (2007). Stereotype threat among school girls in quasi-ordinary classroom circumstances. *Journal of Educational Psychology, 99*, 545-560.

Hyde, Janet S. and Janet E. Mertz, 2009. “Gender, culture and mathematics performance,” PNAS, June 2, vol. 106, no.22, 8801-8807.

Jacobs, Janis E. "Influence of Gender Stereotypes on Parent and Child Mathematics Attitudes." *Journal of Educational Psychology*, 1991, 83(4), pp. 518.

LeFevre, Jo-Anne, Alison G. Kulak, Stephanie L. Heymans, “Factors influencing the selection of university majors varying in mathematical content.” Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement. Vol 24(3), Jul 1992, 276-289.

Murnane, Richard J.; Willett, John B.; Duhaldeborde, Yves and Tyler, John H. "How Important Are the Cognitive Skills of Teenagers in Predicting Subsequent Earnings?" *Journal of Policy Analysis and Management*, 2000, 19(4), pp. 547-68.

Murnane Richard J., John B. Willett, and Frank Levy, “The Growing Importance of Cognitive Skills in Wage Determination,” NBER working paper w5076, March 1995.

Niederle, Muriel, Carmit Segal, and Lise Vesterlund, “How Costly is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness” May 2009.

Niederle, Muriel, and Lise Vesterlund, “Do Women Shy Away from Competition? Do Men Compete too Much?,” *Quarterly Journal of Economics*, August 2007, Vol. 122, No. 3, 1067- 1101.

Niederle, Muriel and Alexandra H. Yestrumskas, “Gender Differences in Seeking Challenges: The Role of Institutions”, NBER working paper, 2008.

Örs, Evren, Frederic Palomino and Eloic Peyrache, “Performance Gender-Gap: Does Competition Matter?”, June 2008, CEPR working paper no. 6891.

Paglin, M. and Rufolo, A.M., 1990. Heterogeneous human capital, occupational choice, and male-female earnings differences. *Journal of Labor Economics* **8**, pp. 123–144.

Pope, Devin and Justin Sydnor, “A New Perspective on Stereotypical Gender Differences in Test Scores”, JEP (this issue).

Preckel, Franzis; Goetz, Thomas; Pekrun, Reinhard and Kleine, Michael. "Gender Differences in Gifted and Average-Ability Students: Comparing Girls' and Boys' Achievement, Self-Concept, Interest, and Motivation in Mathematics." *Gifted Child Quarterly*, 2008, 52(2), pp. 146-59.

Prize, Curtis, 2008a, “Gender, Competition, and Managerial Decisions,” working paper, http://web.usi.edu/crprice/womenmanagers-rnov08.pdf.

Prize, Curtis, 2008b, “Do Women Shy Away From Competition? Do Men Compete Too Much? : A (Failed) Replication,” working paper, http://web.usi.edu/crprice/nvreplication_submit.pdf.

Rothstein, Jesse, 2004, “College Performance Predictions and the SAT”, Journal of Econometrics, volume 121, issue 1-2 (July-August), pp. 297-317.

Segal, Carmit, 2008 “Motivation, Test Scores, and Economic Success” - UPF Working Paper 1124.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women’s math performance. *Journal of Experimental Social Psychology, 35,* 4-28.

Steele, Claude M. (1997). A threat in the air: How stereotypes shape the intellectual identities and performance of women and African Americans. *American Psychologist*, 52, 613-629.

Weinberger, Catherine J. "Mathematical College Majors and the Gender Gap in Wages." *Industrial Relations*, 38(July):407-13, 1999.

Weinberger, Catherine J. "Is Teaching More Girls More Math the Key to Higher Wages?" in Squaring Up: Policy Strategies to Raise Women's Incomes in the U.S., edited by Mary C. King: University of Michigan Press, 2001.

Weinberger, Catherine J. “Is the Science and Engineering Workforce Drawn from the Far Upper Tail of the Math Ability Distribution?”, working paper, 2005.

Wozniak, David, “Choices About Competition: Differences by gender and hormonal fluctuations, and the role of relative performance feedback.” 2009, Working paper, University of Oregon

Xie, Yu and Kimberlee A. Shauman, *Women in Science: Career Processes and Outcomes*, Harvard University, 2003.

Source: Gneezy, Niederle and Rustichini (2003).

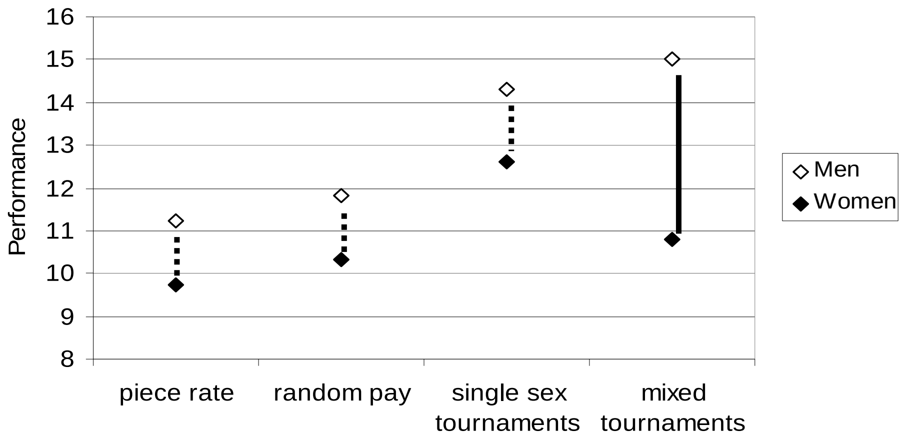

Figure 1
Average Performance of 30 Men and 30 Women in Each Treatment

Source: Niederle and Vesterlund (2007).

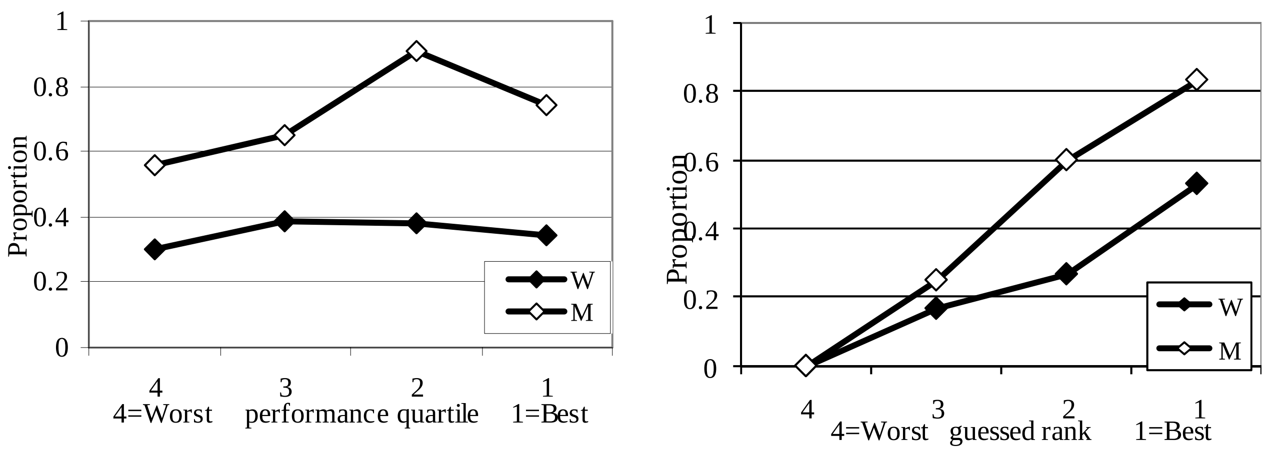

Figure 2: Proportion Selecting Tournament

Panel A: Conditional on Initial Tournament Performance Quartile

Panel B: Conditional on Believed Performance Rank in Initial Tournament