



## **Women in Science and Engineering: What the Research *Really* Says**

A panel discussion co-sponsored by WISELI and the Science Alliance.

Thursday April 14, 2005

Biotechnology Center Auditorium, University of Wisconsin – Madison

“Sex Differences in Math Test Performance: What Do They Mean?”

Prof. Caitlin Allen, Department of Plant Pathology and Women’s Studies Program,  
University of Wisconsin – Madison

That was a fascinating talk. I find myself wondering if [in] my own laboratory if they’re colony mates or competitors. I hope they’re colony mates. Well, I’m going to be talking an area that’s been of interest to me for a while. Though it’s not my immediate research area, I teach a course at the Women’s Study Program that’s called “Biology and Gender” that looks at the evidence for biological bases for cognitive and behavioral differences between the sexes. And one of the kinds of evidence that comes up most frequently in conversations about this is the idea that there’s a difference between men and women in innate ability to do mathematics. Because we have a lot of respect in our culture, particularly in our science culture, for mathematical ability that idea carries a lot of weight and has the potential to do a lot of damage. So it’s been interesting to me to go through the literature on this subject and to try to gather the actual, factual data, some of which I’ll try to present to you today.

So I want to try and tackle three questions. First of all, so that you all know what we're talking about, just how differently do males and females perform on math tests? Then, what kind of evidence suggests or undermines the idea that these differences in test performance are innate, that is you're born with them and they are immutable? And third, why should we care? Does math test performance actually predict success in science and engineering, which is the idea that's implicit in this conversation?

To start with the beginning, I have to say right off the bat that when talk about math tests what we actually mean almost all the time is the Scholastic Aptitude Test. And this is a really powerful dataset. There [are] 1.3 million high school seniors who take the SAT every year and it's one of the major determinants of admission to competitive schools. Paul was talking about the problem of the small N, when we talk about SAT math scores we're actually talking about the problem of the big N. There are so many math scores in this one test that even small differences have a lot of statistical power and therefore give the appearance perhaps of being larger than they actually are.

Nevertheless, over the past thirty-three years, the average male score on the SAT math section has consistently been in the neighborhood of thirty-five points higher than the average female score. And in addition to the difference in the mean, there's also an overrepresentation of males at the very highest end. Nearly twice as many boys as girls score over 700 points on that – you recall that the SAT is graded essentially from 200 to 800 points. And it should be mentioned also, as an aside, that there's greater male

variability at the lower end too. But we don't usually talk about the idiot-gap just usually the genius-gap. It should be said that there's greater male variability in performance on many tests. Females in general have a narrower distribution and it's been suggested that's because women have been socialized to mind better. But I don't know if there's any real evidence for that.

But if you put the distribution curves of male and female scores on top of each other, this is pretty much what they look like [slide depicts distribution]. The female distribution has an average – as I said these are the data for 2004 but they usually look like this. So one of the things you should notice about this if you are scientifically inclined is [that] SAT scores aren't predictive. You couldn't say, "Hmm, I've got somebody with an SAT score of 560, definitely a male." Nor could you say, "This person has a score of 325, must be female." You cannot predict somebody's SAT score, or somebody's sex, by their performance on this test. There's a lot of overlap. And again this reinforces what Paul was saying earlier.

So why is there this difference? Well, there's a fair bit of data on this. First of all, fewer boys than girls take the SAT. About forty-five percent of the SAT test-takers are male and fifty-five percent are female. So there's larger variability toward the lower end of a bunch of indicators in the female sample. And moreover, the males who do take the SAT are a more select group. And what this means is that they have on average compared to the women a higher income of the family, a higher socio-economic status, and a larger proportion of them come from private rather than public high schools. And this is

relevant because more select samples of both males and females do better on all kinds of standardized tests. So private high school students in the U.S. get overall significantly higher scores on the SATs and on many other measures of academic performance. And moreover, the gender gap between males and females in private schools is significantly smaller on the SAT math than the gender gap between males and females in the whole sample. It's only seventeen points as opposed to thirty-five.

A second issue is that the SAT test speed as well as problem solving and if you – in one study anyway, not a large study I should note, but in this one study if you gave the SAT math without a time deadline a difference in performance between boys and girls went away, suggesting that the speediness of the test may have a differential effect across gender.

And then finally, there's the issue of question content. It turns out that math isn't math. If you frame a math question, or for that matter a verbal achievement question, in a familiar or in an alarming or unfamiliar context you can skew the performance of a student on that question. And there's an enormous body of literature on this. It's a well-studied effect that people in education can tell you about all night long. But within the SAT there's some evidence for gender bias as well as racial bias, which I won't talk about tonight but I want to acknowledge. There are more math questions framed in terms of business issues, military problems, sports and fewer framed in terms of relationships, the humanities, the aesthetics, sewing, cooking – subject matter that would be familiar and engaging for female students. So we know that question content is going to effect test

performance and this has been raked over pretty thoroughly by a number of social science researchers.

All right, so what's the evidence that differences in test performance are innate, that is you're born with them? Maybe it isn't innate. Maybe it's the test, because other U.S. tests given on the same population of students don't give the same results as the SAT math, although their N is much smaller. For example, the ACT test, which some students who get into UW-Madison take, has a much smaller gender gap in scores. Another widely used test that's given to all twelfth graders across the U.S., the National Assessment of Educational Progress, finds no gender difference in performance on the math section among U.S. twelfth graders. So maybe there's something special about the SAT. And Janet Hyde in the Psychology Department here has done a nice, large meta-analysis that supports the idea that the SAT is in some ways different from other tests.

On the other hand, there are cultural issues that appear to also effect performance on these kinds of tests. You don't find sex differences in math test performance when you give standardized tests to high school kids in Russia, in India, in Japan, in – Jennifer sent me a paper just today that showed that these differences don't exist in the U.K. Indeed there are big differences across cultures such that girls, and boys too for that matter, in Japan and Singapore score much higher than boys in the U.S. and many other countries. In other words, these tests don't measure something innate. They are measuring something about the training that students have had. In Iceland for whatever reason girls are significantly outscoring the boys on high school math tests. Why Iceland should be so

different from for example Denmark, where they find a reverse gender skew despite cultural similarities, no one has managed to explain it.

What about other ways that we can measure intelligence or cognitive ability? We know that there aren't any sex differences in IQ test scores. We know that males outperform females in some tests of what are called spatial skills. These are a variable set of tests; some of them have to do with rotating an object mentally in three dimensions, that's the most familiar one. However, one of the interesting things about these spatial skills differences is that you can train them away. That if you give these tests to a population of boys and girls, – the study I'm most familiar with is working with eight year-olds – find a gender difference and then you do a three month unit that develops spatial skills, the gender difference goes away and achievement in both groups increases. So that would suggest that this is not an innate but rather a learned set of skills. We also need to note that in terms of grades, which are the currency that most of us are familiar with on college campuses, females on average earn higher grades than males in college and in all majors, from mechanical engineering to art history.

Moreover, and this is of interest to maybe this audience in particular, innate differences aren't compatible with the observed large, recent increases in representation of women in science, math and engineering. So, just to give one example, in 1970 only 0.3 percent of bachelor's degrees in engineering went to women but by 2001 we had somehow become fifty-six times smarter because now we're earning seventeen percent of the engineering

bachelor's degrees. That kind of change is not the result of biology and I have a PhD in Genetics, so I can assure you that it probably was environmental.

But what really matters here, what we should all really be asking, is how useful are these tests? Are they telling us what we want to know about somebody's potential to become a successful engineer and scientist? Larry Summers is by no means alone in his assumption that if you're doing really well on standardized math tests you are headed for the fast-track and a high profile, successful career in science or engineering or math. So for example, one paper remarks, "Individuals with the most potential for mathematics and science are those students who represent the top few centiles in mathematics ability" [Benbow & Arjmand 1990]. Another paper published in *Science* a few years back said, "Substantially fewer females than males score in the upper tails of the math ability distributions," what they really mean here I remind you is the SAT math, "and hence are poised to succeed in the sciences" [Hedges & Newell 1995]. This is one of those things where as the old joke goes, if you can't bother to look up the original citation you say it has long been acknowledged that, in the introduction to your *Science* paper. But it points [to the fact] that this is an assertion, an assumption without much data behind it.

So, in fact is it true? Well, it's not true in college. The stated goal of the SAT is to predict students' first year college grades. But it turns out that when researchers Wainer & Steinberg who I cite here – who actually work for the Educational Testing Service, which is the outfit that runs the SAT – did an analysis, they found that the test was actually not a good predictor of college math course success. And in fact it was a particularly bad

predictor for women. They took a sample size of 47,000 students – not seven, not ten – and found that when they matched women and men who were taking the same kind of course, i.e. remedial math, introductory algebra, trigonometry, calculus, whatever, that women scored about thirty-three points lower on the SAT math before they got into college than men who took the same college math course with these women and earned the same grade. So that means that women went off to college and so did men [and] they took calculus. A certain group of them got As and when you went back and looked at this group that managed to get As in this class you found that the males who actually performed in an equivalent way on this course to the females had indicated much higher potential than the females on the SAT. And this was true not just for people with As but also for Bs, Cs, Ds, and flunking the course. In other words, the SAT math score didn't tell you whether you were likely to succeed in a math course well if you were female.

What about after college? An interesting analysis has been done by an economist at the University of California, Santa Barbara named Catherine Weinberger. And what she did is go looking in the science and engineering workforce and asked what kind of SAT math scores people who were actually out there doing science and engineering jobs had. And first of all, what she found was that fewer than a third of the people in the workforce in science and engineering got SAT math scores over 650. That's a modest SAT math score, that's not the genius-level stuff. And indeed a quarter of the people in the workforce studied, this was not a small study, had actually gotten SAT math scores below 550, which is the average for humanities majors, heaven help us. So, it turns out that you don't



need a really high SAT math score to manage to earn an engineering [degree] and get a job.

“Ah,” you say, what if however they weren’t doing as well in their job? So she did an analysis of how much money they made and she found that having a high SAT math score didn’t affect your earning potential. There was no relationship between your SAT math score in high school and the salary that you were earning however many years later in the workforce.

Okay, so the SAT math is not a particularly good predictor and obviously we’re looking at the mid- to upper range here. We’re talking about 550 and above. It does predict whether somebody is going to go into science or math. If you get a 350 in math on the SAT, chances are you aren’t going to be going into mechanical engineering. But here she’s looking at people who actually finished their degrees. But she also found a curious thing, in terms of pursuit of careers high scoring girls, by which she meant girls who got more than 700 points on the SAT math, were only sixty percent as likely as high scoring boys to pursue SEM careers – science, engineering, math. So these are students who had potential according to their score to go into this field but for some other reason, presumably unrelated to their math score, did not. And so the final conclusion that she came up with was [that] women at all levels of achievement are underrepresented in the SEM workforce relative to men with the same test scores. So the SAT isn’t doing a very good job of predicting success in science and engineering.

And, returning to the social competition question that Linda raised earlier, I was thinking when Lawrence Summers remarks first came out that anybody who has tried to run a research lab and teach at a large academic university knows that math ability is one of many skills that you need to have and that if you're not an efficient multi-tasker, a person who can keep track of your friends and your enemies, a person who can motivate a large group of people to operate, a person who can meet deadlines, who can write a coherent paragraph – I mean the job description is actually rather long and being able to do calculus is not necessarily at the top of the list for being a successful research scientist. But that's an anecdotal perspective.

So to sum up: males do outperform females on some math tests, most notably the SAT math, sex differences in math performance do not appear to be innate and third, math test performance is not a reliable predictor of success in science and engineering coursework or science and engineering careers.