THE SCIENCE OF GENDER AND SCIENCE

Harvard University, April 2005, Psychologists Steven Pinker and Elizabeth Stekle

STEVEN PINKER is the Johnstone Family Professor in the Department of Psychology at Harvard University. His research has won prizes from the National Academy of Sciences and the Royal Institution of Great Britain, and he is the author of six books, including *The Language Instinct*, *How the Mind Works*, *Words and Rules*, and *The Blank Slate*.

ELIZABETH S. SPELKE is Berkman Professor of Psychology at Harvard University, where she is Co-Director of the Mind, Brain, and Behavior Initiative. A member of the National Academy of Sciences and the American Academy of Arts and Sciences, she is cited by *Time* Magazine as one of America's Best in Science and Medicine.

(**STEVEN PINKER:)** Thanks, Liz, for agreeing to this exchange. ...

For those of you who just arrived from Mars, there has been a certain amount of discussion here at Harvard on a particular datum, namely the under-representation of women among tenure-track faculty in elite universities in physical science, math, and engineering.

As with many issues in psychology, there are three broad ways to explain this phenomenon. One can imagine an extreme "nature" position: that males but not females have the talents and temperaments necessary for science. Needless to say, only a madman could take that view. The extreme nature position has no serious proponents.

There is an extreme "nurture" position: that males and females are biologically indistinguishable, and all relevant sex differences are products of socialization and bias.

Then there are various intermediate positions: that the difference is explainable by some combination of biological differences in average temperaments and talents interacting with socialization and bias.

Liz has embraced the extreme nurture position. There is an irony here, because in most discussions in cognitive science she and I are put in the same camp, namely the "innatists," when it comes to explaining the mind. But in this case Liz has said that there is "not a shred of evidence" for the biological factor, that "the evidence against there being an advantage for males in intrinsic aptitude is so overwhelming that it is hard for me to see how one can make a case at this point on the other side," and that "it seems to me as conclusive as any finding I know of in science."

Well we certainly aren't seeing the stereotypical gender difference in confidence here! Now, I'm a controversial guy. I've taken many controversial positions over the

years, and, as a member of Homo sapiens, I think I am right on all of them. But I don't think that in any of them I would say there is "not a shred of evidence" for the other side, even if I think that the evidence favors one side. I would not say that the other side "can't even make a case" for their position, even if I think that their case is not as good as the one I favor. And as for saying that a position is "as conclusive as any finding in science" — well, we're talking about social science here! This statement would imply that the extreme nurture position on gender differences is more conclusive than, say the evidence that the sun is at the center of the solar system, for the laws of thermodynamics, for the theory of evolution, for plate tectonics, and so on.

These are extreme statements — especially in light of the fact that an enormous amount of research, summarized in these and many other literature reviews, in fact points to a very different conclusion. I'll quote from one of them, a book called Sex Differences in Cognitive Ability by Diane Halpern. She is a respected psychologist, recently elected as president of the American Psychological Association, and someone with no theoretical axe to grind. She does not subscribe to any particular theory, and has been a critic, for example, of evolutionary psychology. And here what she wrote in the preface to her book:

"At the time I started writing this book it seemed clear to me that any between sex differences in thinking abilities were due to socialization practices, artifacts, and mistakes in the research. After reviewing a pile of journal articles that stood several feet high, and numerous books and book chapters that dwarfed the stack of journal articles, I changed my mind. The literature on sex differences in cognitive abilities is filled with inconsistent findings, contradictory theories, and emotional claims that are unsupported by the research. Yet despite all the noise in the data, clear and consistent messages could be heard. There are real and in some cases sizable sex differences with respect to some cognitive abilities. Socialization practices are undoubtedly important, but there is also good evidence that biological sex differences play a role in establishing and maintaining cognitive sex differences, a conclusion I wasn't prepared to make when I began reviewing the relevant literature."

This captures my assessment perfectly.

…I want to begin with a confession of my own politics. I am a feminist. …

But it is crucial to distinguish the moral proposition that people should not be discriminated against on account of their sex — which I take to be the core of feminism — and the empirical claim that males and females are biologically indistinguishable. They are not the same thing. Indeed, distinguishing them is essential to protecting the core of feminism. Anyone who takes an honest interest in science has to be prepared for the facts on a given issue to come out either way. And that makes it essential that we not hold the ideals of feminism hostage to the latest findings from the lab or field. Otherwise, if the findings come out as showing a sex difference, one would either have to say, "I guess sex discrimination wasn't so bad

after all," or else furiously suppress or distort the findings so as to preserve the ideal. The truth cannot be sexist. Whatever the facts turn out to be, they should not be taken to compromise the core of feminism.

Why study sex differences? Believe me, being the Bobby Riggs of cognitive science is not my idea of a good time. So should I care about them, especially since they are not the focus of my own research?

First, differences between the sexes are part of the human condition. We all have a mother and a father. Most of us are attracted to members of the opposite sex, and the rest of us notice the difference from those who do. And we can't help but notice the sex of our children, friends, and our colleagues, in every aspect of life.

Also, the topic of possible sex differences is of great scientific interest. Sex is a fundamental problem in biology, and sexual reproduction and sex differences go back a billion years. There's an interesting theory, which I won't have time to explain, which predicts that there should be an overall equal investment of organisms in their sons and daughters; neither sex is predicted to be superior or inferior across the board. There is also an elegant theory, namely Bob Trivers' theory of differential parental investment, which makes highly specific predictions about when you should expect sex differences and what they should look like.

The nature and source of sex differences are also of practical importance. Most of us agree that there are aspects of the world, including gender disparities, that we want to change. But if we want to change the world we must first understand it, and that includes understanding the sources of sex differences.

Let's get back to the datum to be explained. In many ways this is an exotic phenomenon. It involves biologically unprepared talents and temperaments: evolution certainly did not shape any part of the mind to do the work of a professor of mechanical engineering at MIT, for example. The datum has nothing to do with basic cognitive processes, or with those we use in our everyday lives, in school, or even in most college courses, where indeed there are few sex differences.

Also, we are talking about extremes of achievement. Most women are not qualified to be math professors at Harvard because most men aren't qualified to be math professors at Harvard. These are extremes in the population.

And we're talking about a subset of fields. Women are no under-represented to nearly the same extent in all academic fields, and certainly not in all prestigious professions.

Finally, we are talking about a statistical effect. This is such a crucial point that I have to discuss it in some detail.

Women are nowhere near absent even from the field in which they are most under-represented. The explanations for sex differences must be statistical as well. And here is a touchstone for the entire discussion:

These are two Gaussian or normal distributions; two bell curves. The X axis stands for any ability you want to measure. The Yaxis stands for the proportion of people having that ability. The overlapping curves are what you get whenever you compare the sexes on any measure in which they differ. In this example, if we say that this is the male curve and this is the female curve, the means may be different, but at any particular ability level there are always representatives of both genders.

So right away a number of public statements that have been made last couple of months can be seen as red herrings, and should never have been made by anyone who understands the nature of statistical distributions. This includes the accusation that President Summers implied that "50% of the brightest minds in America do not have the right aptitude for science," that "women just can't cut it," and so on. These statements are statistically illiterate, and have nothing to do with the phenomena we are discussing.

There are some important corollaries of having two overlapping normal distributions. One is that a normal distribution falls off according to the negative exponential of the square of the distance from the mean. That means that even when there is only a small difference in the means of two distributions, the more extreme a score, the greater the disparity there will be in the two kinds of individuals having such a score. That is, the ratios get more extreme as you go farther out along the tail. If we hold a magnifying glass to the tail of the distribution, we see that even though the distributions overlap in the bulk of the curves, when you get out to the extremes the difference between the two curves gets larger and larger.

For example, it's obvious that distributions of height for men and women overlap: it's not the case that all men are taller than all women. But while at five foot ten there are thirty men for every woman, at six feet there are two thousand men for every woman. Now, sex differences in cognition tend not to be so extreme, but the statistical phenomenon is the same.

A second important corollary is that tail ratios are affected by differences in variance. And biologists since Darwin have noted that for many traits and many species, males are the more variable gender. So even in cases where the mean for women and the mean for men are the same, the fact that men are more variable implies that the proportion of men would be higher at one tail, and also higher at the other. As it's sometimes summarized: more prodigies, more idiots.

With these statistical points in mind, let me begin the substance of my presentation by connecting the political issue with the scientific one. Economists who study patterns of discrimination have long argued (generally to no avail) that there is a crucial conceptual difference between difference and discrimination. A departure from a 50-50 sex ratio in any profession does not, by itself, imply that we are seeing discrimination, unless the interests and aptitudes of the two groups are equated. Let me illustrate the point with an example, involving myself.

I work in a scientific field — the study of language acquisition in children — that is in fact dominated by women. Seventy-five percent of the members the main professional association are female, as are a majority of the keynote speakers at our main conference. I'm here to tell you that it's not because men like me have been discriminated against. I decided to study language development, as opposed to, say, mechanical engineering, for many reasons. The goal of designing a better automobile transmission does not turn me on as much as the goal of figuring out how kids acquire language. And I don't think I'd be as good at designing a transmission as I am in studying child language.

Now, all we need to do to explain sex differences without invoking the discrimination or invidious sexist comparisons is to suppose that whatever traits I have that predispose me to choose (say) child language over (say) mechanical engineering are not exactly equally distributed statistically among men and women. For those of you out there — of either gender — who also are not mechanical engineers, you should understand what I'm talking about.

Okay, so what are the similarities and differences between the sexes? There certainly are many similarities. Men and women show no differences in general intelligence or g — on average, they are exactly the same, right on the money. Also, when it comes to the basic categories of cognition — how we negotiate the world and live our lives; our concept of objects, of numbers, of people, of living things, and so on — there are no differences.

Indeed, in cases where there are differences, there are as many instances in which women do slightly better than men as ones in which men do slightly better than women. For example, men are better at throwing, but women are more dexterous. Men are better at mentally rotating shapes; women are better at visual memory. Men are better at mathematical problem-solving; women are better at mathematical calculation. And so on.

But there are at least six differences that are relevant to the datum we have been discussing. The literature on these differences is so enormous that I can only touch on a fraction of it. I'll restrict my discussion to a few examples in which there are enormous data sets, or there are meta-analyses that boil down a literature.

The first difference, long noted by economists studying employment practices, is that men and women differ in what they state are their priorities in life. To sum it up: men, on average, are more likely to chase status at the expense of their families; women give a more balanced weighting. Once again: Think statistics! The finding is not that women value family and don't value status. It is not that men value status and don't value family. Nor does the finding imply that every last woman has the asymmetry that women show on average or that every last man has the asymmetry that men show on average. But in large data sets, on average, an asymmetry what you find.

Just one example. In a famous long-term study of mathematically precocious youth, 1,975 youngsters were selected in 7th grade for being in the top 1% of ability in mathematics, and then followed up for more than two decades. These men and women are certainly equally talented. And if anyone has ever been encouraged in math and science, these kids were. Both genders: they are equal in their levels of achievement, and they report being equally satisfied with the course of their lives. Nonetheless there are statistical differences in what they say is important to them. There are some things in life that the females rated higher than males, such as the ability to have a part-time career for a limited time in one's life; living close to parents and relatives; having a meaningful spiritual life; and having strong friendships. And there are some things in life that the males rated higher than the females. They include having lots of money; inventing or creating something; having a full-time career; and being successful in one's line of work. It's worth noting that studies of highly successful people find that single-mindedness and competitiveness are recurring traits in geniuses (of both sexes).

Here is one other figure from this data set. As you might expect, this sample has a lot of people who like to work Herculean hours. Many people in this group say they would like to work 50, 60, even 70 hours a week. But there are also slight differences. At each one of these high numbers of hours there are slightly more men than women who want to work that much. That is, more men than women don't care about whether they have a life.

Second, interest in people versus things and abstract rule systems. There is a staggering amount of data on this trait, because there is an entire field that studies people's vocational interests. I bet most of the people in this room have taken a vocational interest test at some point in their lives. And this field has documented that there are consistent differences in the kinds of activities that appeal to men and women in their ideal jobs. I'll just discuss one of them: the desire to work with people versus things. There is an enormous average difference between women and men in this dimension, about one standard deviation.

And this difference in interests will tend to cause people to gravitate in slightly different directions in their choice of career. The occupation that fits best with the "people" end of the continuum is "director of a community services organization." The occupations that fit best with the "things" end are physicist, chemist, mathematician, computer programmer, and biologist.

We see this consequence not only in the choice of whether to go into science, but also in the choice which branch of science the two sexes tend to go into. Needless to say, from 1970 to 2002 there was a huge increase in the percentage of university degrees awarded to women. But the percentage still differs dramatically across fields. Among the Ph.Ds awarded in 2001, for example, in education 65% of the doctorates went to women; in the social sciences, 54%; in the life sciences, 47%; in the physical sciences, 26%; in engineering, 17%. This is completely predictable from the difference in interests between people and living things, on the one hand, and

inanimate objects, on the other. And the pattern is pretty much the same in 1980 and 2001, despite the change in absolute numbers.

Third, risk. Men are by far the more reckless sex. In a large meta-analysis involving 150 studies and 100,000 participants, in 14 out of 16 categories of risk-taking, men were over-represented. The two sexes were equally represented in the other two categories, one of which was smoking, for obvious reasons. And two of the largest sex differences were in "intellectual risk taking" and "participation in a risky experiment." We see this sex difference in everyday life, in particular, in the following category: the Darwin Awards, "commemorating those individuals who ensure the long-term survival of our species by removing themselves from the gene pool in a sublimely idiotic fashion." Virtually all — perhaps all — of the winners are men.

Fourth, three-dimensional mental transformations: the ability to determine whether the drawings in each of these pairs the same 3-dimensional shape. Again I'll appeal to a meta-analysis, this one containing 286 data sets and 100,000 subjects. The authors conclude, "we have specified a number of tests that show highly significant sex differences that are stable across age, at least after puberty, and have not decreased in recent years." Now, as I mentioned, for some kinds of spatial ability, the advantage goes to women, but in "mental rotation,"spatial perception," and "spatial visualization" the advantage goes to men.

Now, does this have any relevance to scientific achievement? We don't know for sure, but there's some reason to think that it does. In psychometric studies, three-dimensional spatial visualization is correlated with mathematical problem-solving. And mental manipulation of objects in three dimensions figures prominently in the memoirs and introspections of most creative physicists and chemists, including Faraday, Maxwell, Tesla, Kéekulé, and Lawrence, all of whom claim to have hit upon their discoveries by dynamic visual imagery and only later set them down in equations. A typical introspection is the following: "The cyclical entities which seem to serve as elements in my thought are certain signs and more or less clear images which can be voluntarily reproduced and combined. This combinatory play seems to be the essential feature in productive thought before there is any connection with logical construction in words or other kinds of signs." The quote comes from this fairly well-known physicist.

Fifth, mathematical reasoning. Girls and women get better school grades in mathematics and pretty much everything else these days. And women are better at mathematical calculation. But consistently, men score better on mathematical word problems and on tests of mathematical reasoning, at least statistically. Again, here is a meta analysis, with 254 data sets and 3 million subjects. It shows no significant difference in childhood; this is a difference that emerges around puberty, like many secondary sexual characteristics. But there are sizable differences in adolescence and adulthood, especially in high-end samples. Here is an example of the average SAT mathematical scores, showing a 40-point difference in favor of men that's pretty much consistent from 1972 to 1997. In the Study of Mathematically

Precocious Youth (in which 7th graders were given the SAT, which of course ordinarily is administered only to older, college-bound kids), the ratio of those scoring over 700 is 2.8 to 1 male to female. (Admittedly, and interestingly, that's down from 25 years ago, when the ratio was 13-to1, and perhaps we can discuss some of the reasons.) At the 760 cutoff, the ratio nowadays is 7 males to 1 female.

Now why is there a discrepancy with grades? Do SATs and other tests of mathematical reasoning aptitude underpredict grades, or do grades overpredict high-end aptitude? At the Radical Forum Liz was completely explicit in which side she takes, saying that "the tests are no good," unquote. But if the tests are really so useless, why does every major graduate program in science still use them — including the very departments at Harvard and MIT in which Liz and I have selected our own graduate students?

I think the reason is that school grades are affected by homework and by the ability to solve the kinds of problems that have already been presented in lecture and textbooks. Whereas the aptitude tests are designed to test the application of mathematical knowledge to unfamiliar problems. And this, of course, is closer to the way that math is used in actually doing math and science.

Indeed, contrary to Liz, and the popular opinion of many intellectuals, the tests are surprisingly good. There is an enormous amount of data on the predictive power of the SAT. For example, people in science careers overwhelmingly scored in 90th percentile in the SAT or GRE math test. And the tests predict earnings, occupational choice, doctoral degrees, the prestige of one's degree, the probability of having a tenure-track position, and the number of patents. Moreover this predictive power is the same for men and for women. As for why there is that underprediction of grades — a slight under-prediction, one-tenth of a standard deviation — the Educational Testing Service did a study on that phenomenon, and were able to explain the mystery by a combination of the choice of major, which differs between the sexes, and the greater conscientiousness of women.

Finally there's a sex difference in variability. It's crucial here to look at the right samples. Estimates of variance depend highly on the tails of the distribution, which by definition contain smaller numbers of people. Since people at the tails of the distribution in many surveys are likely to be weeded out for various reasons, it's important to have large representative samples from national populations. In this regard the gold standard is the Science paper by Novell and Hedges, which reported six large stratified probability samples. They found that in 35 out of 37 tests, including all of the tests in math, space, and science, the male variance was greater than the female variance.

One other data set meeting the gold standard is displayed in this graph, showing the entire population of Scotland, who all took an intelligence test in a single year. The X axis represents IQ, where the mean is 100, and the Yaxis represents the proportion of men versus women. As you can see these are extremely orderly data. In the middle part of the range, females predominate; at both

extremes, males slightly predominate. Needless to say, there is a large percentage of women at both ends of the scale — but there is also large sex difference.

Now the fact that these six gender differences exist does not mean that they are innate. This of course is a much more difficult issue to resolve. A necessary preamble to this discussion is that nature and nurture are not alternatives; it is possible that the explanation for a given sex difference involves some of each. The only issue is whether the contribution of biology is greater than zero. I think that there are ten kinds of evidence that the contribution of biology is greater than zero, though of course it is nowhere near 100 percent.

First, there are many biological mechanisms by which a sex difference could occur. There are large differences between males and females in levels of sex hormones, especially prenatally, in the first six months of life, and in adolescence. There are receptors for hormones all over the brain, including the cerebral cortex. There are many small differences in men's and women's brains, including the overall size of the brain (even correcting for body size), the density of cortical neurons, the degree of cortical asymmetry, the size of hypothalamic nuclei, and several others.

Second, many of the major sex differences — certainly some of them, maybe all of them, are universal. The idea that there are cultures out there somewhere in which everything is the reverse of here turns out to be an academic legend. In his survey of the anthropological literature called Human Universals, the anthropologist Donald Brown points out that in all cultures men and women are seen as having different natures; that there is a greater involvement of women in direct child care; more competitiveness in various measures for men than for women; and a greater spatial range traveled by men compared to by women.

In personality, we have a cross-national survey (if not a true cross-cultural one) in Feingold's meta-analysis, which noted that gender differences in personality are consistent across ages, years of data collection, educational levels, and nations. When it comes to spatial manipulation and mathematical reasoning, we have fewer relevant data, and we honestly don't have true cross-cultural surveys, but we do have cross-national surveys. David Geary and Catherine Desoto found the expected sex difference in mental rotation in ten European countries and in Ghana, Turkey, and China. Similarly, Diane Halpern, analyzing results from ten countries, said that "the majority of the findings show amazing cross-cultural consistency when comparing males and females on cognitive tests."

Third, stability over time. Surveys of life interests and personality have shown little or no change in the two generations that have come of age since the second wave of feminism. There is also, famously, resistance to change in communities that, for various ideological reasons, were dedicated to stamping out sex differences, and found they were unable to do so. These include the Israeli kibbutz, various American Utopian communes a century ago, and contemporary androgynous academic couples.

In tests of mental rotation, the meta-analysis by Voyer et al found no change over time. In mathematical reasoning there has been a decline in the size of the difference, although it has certainly not disappeared.

Fourth, many sex differences can be seen in other mammals. It would be an amazing coincidence if these differences just happened to be replicated in the arbitrary choices made by human cultures at the dawn of time. There are large differences between males and females in many mammals in aggression, in investment in offspring, in play aggression play versus play parenting, and in the range size, which predicts a species' sex differences in spatial ability (such as in solving mazes), at least in polygynous species, which is how the human species is classified. Many primate species even show a sex difference in their interest in physical objects versus conspecifics, a difference seen their patterns of juvenile play. Among baby vervet monkeys, the males even prefer to play with trucks and the females with other kinds of toys!

Fifth, many of these differences emerge in early childhood. It is said that there is a technical term for people who believe that little boys and little girls are born indistinguishable and are molded into their natures by parental socialization. The term is "childless."

Some sex differences seem to emerge even in the first week of life. Girls respond more to sounds of distress, and girls make more eye contact than boys. And in a study that I know Liz disputes and that I hope we'll talk about, newborn boys were shown to be more interested in looking at a physical object than a face, whereas newborn girls were shown to be more interested in looking at a face than a physical object.

A bit later in development there are vast and robust differences between boys and girls, seen all over the world. Boys far more often than girls engage in rough-and-tumble play, which involves aggression, physical activity, and competition. Girls spend a lot more often in cooperative play. Girls engage much more often in play parenting. And yes, boys the world over turn anything into a vehicle or a weapon, and girls turn anything into a doll. There are sex differences in intuitive psychology, that is, how well children can read one another's minds. For instance, several large studies show that girls are better than boys in solving the "false belief task," and in interpreting the mental states of characters in stories.

Sixth, genetic boys brought up as girls. In a famous 1970s incident called the John/Joan case, one member of a pair of identical twin boys lost his penis in a botched circumcision (I was relieved to learn that this was not done by a moyl, but by a bumbling surgeon). Following advice from the leading gender expert of the time, the parents agreed to have the boy castrated, given female-specific hormones, and brought up as a girl. All this was hidden from him throughout his childhood.

When I was an undergraduate the case was taught to me as proof of how gender roles are socially acquired. But it turned out that the facts had been suppressed.

When "Joan" and her family were interviewed years later, it turned out that from the youngest ages he exhibited boy-typical patterns of aggression and rough-and-tumble play, rejected girl-typical activities, and showed a greater interest in things than in people. At age 14, suffering from depression, his father finally told him the truth. He underwent further surgery, married a woman, adopted two children, and got a job in a slaughterhouse.

This is not just a unique instance. In a condition called cloacal exstrophy, genetic boys are sometimes born without normal male genitalia. When they are castrated and brought up as girls, in 25 out of 25 documented instances they have felt that they were boys trapped in girls' bodies, and showed male-specific patterns of behavior such as rough-and-tumble play.

Seventh, a lack of differential treatment by parents and teachers. These conclusions come as a shock to many people. One comes from Lytton and Romney's meta-analysis of sex-specific socialization involving 172 studies and 28,000 children, in which they looked both at parents' reports and at direct observations of how parents treat their sons and daughters — and found few or no differences among contemporary Americans. In particular, there was no difference in the categories "Encouraging Achievement" and "Encouraging Achievement in Mathematics."

There is a widespread myth that teachers (who of course are disproportionately female) are dupes who perpetuate gender inequities by failing to call on girls in class, and who otherwise having low expectations of girls' performance. In fact Jussim and Eccles, in a study of 100 teachers and 1,800 students, concluded that teachers seemed to be basing their perceptions of students on those students' actual performances and motivation.

Eighth, studies of prenatal sex hormones: the mechanism that makes boys boys and girls girls in the first place. There is evidence, admittedly squishy in parts, that differences in prenatal hormones make a difference in later thought and behavior even within a given sex. In the condition called congenital adrenal hyperplasia, girls in utero are subjected to an increased dose of androgens, which is neutralized postnatally. But when they grow up they have male-typical toy preferences — trucks and guns — compared to other girls, male-typical play patterns, more competitiveness, less cooperativeness, and male-typical occupational preferences. However, research on their spatial abilities is inconclusive, and I cannot honestly say that there are replicable demonstrations that CAH women have male-typical patterns of spatial cognition.

Similarly, variations in fetal testosterone, studied in various ways, show that fetal testosterone has a nonmonotic relationship to reduced eye contact and face perception at 12 months, to reduced vocabulary at 18 months, to reduced social skills and greater narrowness of interest at 48 months, and to enhanced mental rotation abilities in the school-age years.

Ninth, circulating sex hormones. I'm going to go over this slide pretty quickly because the literature is a bit messy. Though it's possible that all claims of the effects of hormones on cognition will turn out to be bogus, I suspect something will be salvaged from this somewhat contradictory literature. There are, in any case, many studies showing that testosterone levels in the low-normal male range are associated with better abilities in spatial manipulation. And in a variety of studies in which estrogens are compared or manipulated, there is evidence, admittedly disputed, for statistical changes in the strengths and weaknesses in women's cognition during the menstrual cycle, possibly a counterpart to the changes in men's abilities during their daily and seasonal cycles of testosterone.

My last kind of evidence: imprinted X chromosomes. In the past fifteen years an entirely separate genetic system capable of implementing sex differences has been discovered. In the phenomenon called genetic imprinting, studied by David Haig and others, a chromosome such as the X chromosome can be altered depending on whether it was passed on from one's mother or from one's father. This makes a difference in the condition called Turner syndrome, in which a child has just one X chromosome, but can get it either from her mother or her father. When she inherits an X that is specific to girls, on average she has a better vocabulary and better social skills, and is better at reading emotions, at reading body language, and at reading faces.

…Are these stereotypes? Yes, many of them are (although, I must add, not all of them — for example, women's superiority in spatial memory and mathematical calculation. There seems to be a widespread assumption that if a sex difference conforms to a stereotype, the difference must have been caused by the stereotype, via differential expectations for boys and for girls. But of course the causal arrow could go in either direction: stereotypes might reflect differences rather than cause them. In fact there's an enormous literature in cognitive psychology which says that people can be good intuitive statisticians when forming categories and that their prototypes for conceptual categories track the statistics of the natural world pretty well. For example, there is a stereotype that basketball players are taller on average than jockeys. But that does not mean that basketball players grow tall, and jockeys shrink, because we expect them to have certain heights! Likewise, Alice Eagly and Jussim and Eccles have shown that most of people's gender stereotypes are in fact pretty accurate. Indeed the error people make is in the direction of underpredicting sex differences.

To sum up: I think there is more than "a shred of evidence" for sex differences that are relevant to statistical gender disparities in elite hard science departments. There are reliable average difference in life priorities, in an interest in people versus things, in risk-seeking, in spatial transformations, in mathematical reasoning, and in variability in these traits. And there are ten kinds of evidence that these differences are not completely explained by socialization and bias, although they surely are in part.

A concluding remark. None of this provides grounds for ignoring the biases and barriers that do keep women out of science, as long as we keep in mind the distinction between fairness on the one hand and sameness on the other. And I will give the final word to Gloria Steinem: "there are very few jobs that actually require a penis or a vagina, and all the other jobs should be open to both sexes."

(**ELIZABETH SPELKE**:) Thanks, especially to Steve; I'm really glad we're able to have this debate, I've been looking forward to it.

I want to start by talking about the points of agreement between Steve and me, and as he suggested, there are many. If we got away from the topic of sex and science, we'd be hard pressed to find issues that we disagree on. Here are a few of the points of agreement that are particularly relevant to the discussions of the last few months.

First, we agree that both our society in general and our university in particular will be healthiest if all opinions can be put on the table and debated on their merits. We also agree that claims concerning sex differences are empirical, they should be evaluated by evidence, and we'll all be happier and live longer if we can undertake that evaluation as dispassionately and rationally as possible. We agree that the mind is not a blank slate; in fact one of the deepest things that Steve and I agree on is that there is such a thing as human nature, and it is a fascinating and exhilarating experience to study it. And finally, I think we agree that the role of scientists in society is rather modest. Scientists find things out. The much more difficult questions of how to use that information, live our lives, and structure our societies are not questions that science can answer. Those are questions that everybody must consider.

So where do we disagree?

We disagree on the answer to the question, why in the world are women scarce as hens' teeth on Harvard's mathematics faculty and other similar institutions? In the current debate, two classes of factors have been said to account for this difference. In one class are social forces, including overt and covert discrimination and social influences that lead men and women to develop different skills and different priorities. In the other class are genetic differences that predispose men and women to have different capacities and to want different things.

In his book, The Blank Slate, and again today, Steve argued that social forces are over-rated as causes of gender differences. Intrinsic differences in aptitude are a larger factor, and intrinsic differences in motives are the biggest factor of all. Most of the examples that Steve gave concerned what he takes to be biologically based differences in motives.

My own view is different. I think the big forces causing this gap are social factors. There are no differences in overall intrinsic aptitude for science and mathematics between women and men. Notice that I am not saying the genders are indistinguishable, that men and women are alike in every way, or even that men and women have identical cognitive profiles. I'm saying that when you add up all the

things that men are good at, and all the things that women are good at, there is no overall advantage for men that would put them at the top of the fields of math and science.

On the issue of motives, I think we're not in a position to know whether the different things that men and women often say they want stem only from social forces, or in part from intrinsic sex differences. I don't think we can know that now.

I want to start with the issue that's clearly the biggest source of debate between Steve and me: the issue of differences in intrinsic aptitude. This is the only issue that my own work and professional knowledge bear on. Then I will turn to the social forces, as a lay person as it were, because I think they are exerting the biggest effects. Finally, I'll consider the question of intrinsic motives, which I hope we'll come back to in our discussion.

Over the last months, we've heard three arguments that men have greater cognitive aptitude for science. The first argument is that from birth, boys are interested in objects and mechanics, and girls are interested in people and emotions. The predisposition to figure out the mechanics of the world sets boys on a path that makes them more likely to become scientists or mathematicians. The second argument assumes, as Galileo told us, that science is conducted in the language of mathematics. On the second claim, males are intrinsically better at mathematical reasoning, including spatial reasoning. The third argument is that men show greater variability than women, and as a result there are more men at the extreme upper end of the ability distribution from which scientists and mathematicians are drawn. Let me take these claims one by one.

The first claim, as Steve said, is gaining new currency from the work of Simon Baron-Cohen. It's an old idea, presented with some new language. Baron-Cohen says that males are innately predisposed to learn about objects and mechanical relationships, and this sets them on a path to becoming what he calls "systematizers." Females, on the other hand, are innately predisposed to learn about people and their emotions, and this puts them on a path to becoming "empathizers." Since systematizing is at the heart of math and science, boys are more apt to develop the knowledge and skills that lead to math and science.

To anyone as old as I am who has been following the literature on sex differences, this may seem like a surprising claim. The classic reference on the nature and development of sex differences is a book by Eleanor Maccoby and Carol Jacklin that came out in the 1970s. They reviewed evidence for all sorts of sex differences, across large numbers of studies, but they also concluded that certain ideas about differences between the genders were myths. At the top of their list of myths was the idea that males are primarily interested in objects and females are primarily interested in people. They reviewed an enormous literature, in which babies were presented with objects and people to see if they were more interested in one than the other. They concluded that there were no sex differences in these interests.

Nevertheless, this conclusion was made in the early 70s. At that time, we didn't know much about babies' understanding of objects and people, or how their understanding grows. Since Baron-Cohen's claims concern differential predispositions to learn about different kinds of things, you could argue that the claims hadn't been tested in Maccoby and Jacklin's time. What does research now show?

Let me take you on a whirlwind tour of 30 years of research in one powerpoint slide. From birth, babies perceive objects. They know where one object ends and the next one begins. They can't see objects as well as we can, but as they grow their object perception becomes richer and more differentiated.

Babies also start with rudimentary abilities to represent that an object continues to exist when it's out of view, and they hold onto those representations longer, and over more complicated kinds of changes, as they grow. Babies make basic inferences about object motion: inferences like, the force with which an object is hit determines the speed with which it moves. These inferences undergo regular developmental changes over the infancy period.

In each of these cases, there is systematic developmental change, and there's variability. Because of this variability, we can compare the abilities of male infants to females. Do we see sex differences? The research gives a clear answer to this question: We don't.

Male and female infants are equally interested in objects. Male and female infants make the same inferences about object motion, at the same time in development. They learn the same things about object mechanics at the same time.

Across large numbers of studies, occasionally a study will favor one sex over the other. For example, girls learn that the force with which something is hit influences the distance it moves a month earlier than boys do. But these differences are small and scattered. For the most part, we see high convergence across the sexes. Common paths of learning continue through the preschool years, as kids start manipulating objects to see if they can get a rectangular block into a circular hole. If you look at the rates at which boys and girls figure these things out, you don't find any differences. We see equal developmental paths.

I think this research supports an important conclusion. In discussions of sex differences, we need to ask what's common across the two sexes. One thing that's common is infants don't divide up the labor of understanding the world, with males focusing on mechanics and females focusing on emotions. Male and female infants are both interested in objects and in people, and they learn about both. The conclusions that Maccoby and Jacklin drew in the early 1970s are well supported by research since that time.

Let me turn to the second claim. People may have equal abilities to develop intuitive understanding of the physical world, but formal math and science don't build on these intuitions. Scientists use mathematics to come up with new

characterizations of the world and new principles to explain its functioning. Maybe males have an edge in scientific reasoning because of their greater talent for mathematics.

As Steve said, formal mathematics is not something we have evolved to do; it's a recent accomplishment. Animals don't do formal math or science, and neither did humans back in the Pleistocene. If there is a biological basis for our mathematical reasoning abilities, it must depend on systems that evolved for other purposes, but that we've been able to harness for the new purpose of representing and manipulating numbers and geometry.

Research from the intersecting fields of cognitive neuroscience, neuropsychology, cognitive psychology, and cognitive development provide evidence for five "core systems" at the foundations of mathematical reasoning. The first is a system for representing small exact numbers of objects — the difference between one, two, and three. This system emerges in human infants at about five months of age, and it continues to be present in adults. The second is a system for discriminating large, approximate numerical magnitudes — the difference between a set of about ten things and a set of about 20 things. That system also emerges early in infancy, at four or five months, and continues to be present and functional in adults.

The third system is probably the first uniquely human foundation for numerical abilities: the system of natural number concepts that we construct as children when we learn verbal counting. That construction takes place between about the ages of two and a half and four years. The last two systems are first seen in children when they navigate. One system represents the geometry of the surrounding layout. The other system represents landmark objects.

All five systems have been studied quite extensively in large numbers of male and female infants. We can ask, are there sex differences in the development of any of these systems at the foundations of mathematical thinking? Again, the answer is no. I will show you data from just two cases.

The first is the development of natural number concepts, constructed by children between the ages of two and four. At any particular time in this period, you'll find a lot of variability. For example, between the ages of three and three and a half years, some children have only figured out the meaning of the word "one" and can only distinguish the symbolic concept one from all other numbers. Other kids have figured out the meanings of all the words in the count list up to "ten" or more, and they can use all of them in a meaningful way. Most kids are somewhere in between: they have figured out the first two symbols, or the first three, and so forth. When you compare children's performance by sex, you see no hint of a superiority of males in constructing natural number concepts.

The other example comes from studies that I think are the closest thing in preschool children to the mental rotation tests conducted with adults. In these studies, children are brought into a room of a given shape, something is hidden in a

corner, and then their eyes are closed and they're spun around. They have to remember the shape of the room, open their eyes, and figure out how to rotate themselves back to the object where it was hidden. If you test a group of 4 year olds, you find they can do this task well above chance but not perfectly; there's a range of performance. When you break that performance down by gender, again there is not a hint of an advantage for boys over girls.

These findings and others support two important points. First, indeed there is a biological foundation to mathematical and scientific reasoning. We are endowed with core knowledge systems that emerge prior to any formal instruction and that serve as a basis for mathematical thinking. Second, these systems develop equally in males and females. Ten years ago, the evolutionary psychologist and sex difference researcher, David Geary, reviewed the literature that was available at that time. He concluded that there were no sex differences in "primary abilities" underlying mathematics. What we've learned in the last ten years continues to support that conclusion.

Sex differences do emerge at older ages. Because they emerge later in childhood, it's hard to tease apart their biological and social sources. But before we attempt that task, let's ask what the differences are.

I think the following is a fair statement, both of the cognitive differences that Steve described and of others. When people are presented with a complex task that can be solved through multiple different strategies, males and females sometimes differ in the strategy that they prefer.

For example, if a task can only be solved by representing the geometry of the layout, we do not see a difference between men and women. But if the task can be accomplished either by representing geometry or by representing individual landmarks, girls tend to rely on the landmarks, and boys on the geometry. To take another example, when you compare the shapes of two objects of different orientations, there are two different strategies you can use. You can attempt a holistic rotation of one of the objects into registration with the other, or you can do point-by-point featural comparisons of the two objects. Men are more likely to do the first; women are more likely to do the second.

Finally, the mathematical word problems on the SAT-M very often allow multiple solutions. Both item analyses and studies of high school students engaged in the act of solving such problems suggest that when students have the choice of solving a problem by plugging in a formula or by doing Ven diagram-like spatial reasoning, girls tend to do the first and boys tend to do the second.

Because of these differences, males and females sometimes show differing cognitive profiles on timed tests. When you have to solve problems fast, some strategies will be faster than others. Thus, females perform better at some verbal, mathematical and spatial tasks, and males perform better at other verbal, mathematical, and spatial tasks. This pattern of differing profiles is not well captured by the

generalization, often bandied about in the popular press, that women are "verbal" and men are "spatial." There doesn't seem to be any more evidence for that than there was for the idea that women are people-oriented and men are object-oriented. Rather the differences are more subtle.

Does one of these two profiles foster better learning of math than the other? In particular, is the male profile better suited to high-level mathematical reasoning.

At this point, we face a question that's been much discussed in the literature on mathematics education and mathematical testing. The question is, by what yardstick can we decide whether men or women are better at math?

Some people suggest that we look at performance on the SAT-M, the quantitative portion of the Scholastic Assessment Test. But this suggestion raises a problem of circularity. The SAT test is composed of many different types of items. Some of those items are solved better by females. Some are solved better by males. The people who make the test have to decide, how many items of each type to include? Depending on how they answer that question, they can create a test that makes women look like better mathematicians, or a test that makes men look like better mathematicians. What's the right solution?

Books are devoted to this question, with much debate, but there seems to be a consensus on one point: The only way to come up with a test that's fair is to develop an independent understanding of what mathematical aptitude is and how it's distributed between men and women. But in that case, we can't use performance on the SAT to give us that understanding. We've got to get that understanding in some other way. So how are we going to get it?

A second strategy is to look at job outcomes. Maybe the people who are better at mathematics are those who pursue more mathematically intensive careers. But this strategy raises two problems. First, which mathematically intensive jobs should we choose? If we choose engineering, we will conclude that men are better at math because more men become engineers. If we choose accounting, we will think that women are better at math because more women become accountants: 57% of current accountants are women. So which job are we going to pick, to decide who has more mathematical talent?

These two examples suggest a deeper problem with job outcomes as a measure of mathematical talent. Surely you've got to be good at math to land a mathematically intensive job, but talent in mathematics is only one of the factors influencing career choice. It can't be our gold standard for mathematical ability.

So what can be? I suggest the following experiment. We should take a large number of male students and a large number of female students who have equal educational backgrounds, and present them with the kinds of tasks that real mathematicians face. We should give them new mathematical material that they have not yet mastered, and allow them to learn it over an extended period of time: the kind of time scale that real mathematicians work on. We should ask, how well do the

students master this material? The good news is, this experiment is done all the time. It's called high school and college.

Here's the outcome. In high school, girls and boys now take equally many math classes, including the most advanced ones, and girls get better grades. In college, women earn almost half of the bachelor's degrees in mathematics, and men and women get equal grades. Here I respectfully disagree with one thing that Steve said: men and women get equal grades, even when you only compare people within a single institution and a single math class. Equating for classes, men and women get equal grades.

The outcome of this large-scale experiment gives us every reason to conclude that men and women have equal talent for mathematics. Here, I too would like to quote Diane Halpern. Halpern reviews much evidence for sex differences, but she concludes, "differences are not deficiencies." Men and women have equal aptitude for mathematics. Yes, there are sex differences, but they don't add up to an overall advantage for one sex over the other.

Let me turn to the third claim, that men show greater variability, either in general or in quantitative abilities in particular, and so there are more men at the upper end of the ability distribution. I can go quickly here, because Steve has already talked about the work of Camilla Benbow and Julian Stanley, focusing on mathematically precocious youth who are screened at the age of 13, put in intensive accelerated programs, and then followed up to see what they achieve in mathematics and other fields.

As Steve said, students were screened at age 13 by the SAT, and there were many more boys than girls who scored at the highest levels on the SAT-M. In the 1980s, the disparity was almost 13 to 1. It is now substantially lower, but there still are more boys among the very small subset of people from this large, talented sample who scored at the very upper end. Based on these data, Benbow and Stanley concluded that there are more boys than girls in the pool from which future mathematicians will be drawn. But notice the problem with this conclusion: It's based entirely on the SAT-M. This test, and the disparity it revealed, are in need of an explanation, a firmer yardstick for assessing and understanding gender differences in this talented population.

Fortunately, Benbow, Stanley and Lubinski have collected much more data on these mathematically talented boys and girls: not just the ones with top scores on one timed test, but rather the larger sample of girls and boys who were accelerated and followed over time. Let's look at some of the key things that they found.

First, they looked at college performance by the talented sample. They found that the males and females took equally demanding math classes and majored in math in equal numbers. More girls majored in biology and more boys in physics and engineering, but equal numbers of girls and boys majored in math. And they got equal grades. The SAT-M not only under-predicts the performance of college women

in general, it also under-predicted the college performance of women in the talented sample. These women and men have been shown to be equally talented by the most meaningful measure we have: their ability to assimilate new, challenging material in demanding mathematics classes at top-flight institutions. By that measure, the study does not find any difference between highly talented girls and boys.

So, what's causing the gender imbalance on faculties of math and science? Not differences in intrinsic aptitude. Let's turn to the social factors that I think are much more important. Because I'm venturing outside my own area of work, and because time is short, I won't review all of the social factors producing differential success of men and women. I will talk about just one effect: how gender stereotypes influence the ways in which males and females are perceived.

Let me start with studies of parents' perceptions of their own children. Steve said that parents report that they treat their children equally. They treat their boys and girls alike, and they encourage them to equal extents, for they want both their sons and their daughters to succeed. This is no doubt true. But how are parents perceiving their kids?

Some studies have interviewed parents just after the birth of their child, at the point where the first question that 80% of parents ask — is it a boy or a girl? — has been answered. Parents of boys describe their babies as stronger, heartier, and bigger than parents of girls. The investigators also looked at the babies' medical records and asked whether there really were differences between the boys and girls in weight, strength, or coordination. The boys and girls were indistinguishable in these respects, but the parents' descriptions were different.

At 12 months of age, girls and boys show equal abilities to walk, crawl, or clamber. But before one study, Karen Adolph, an investigator of infants' locomotor development, asked parents to predict how well their child would do on a set of crawling tasks: Would the child be able to crawl down a sloping ramp? Parents of sons were more confident that their child would make it down the ramp than parents of daughters. When Adolph tested the infants on the ramp, there was no difference whatever between the sons and daughters, but there was a difference in the parents' predictions.

My third example, moving up in age, comes from the studies of Jackie Eccles. She asked parents of boys and girls in sixth grade, how talented do you think your child is in mathematics? Parents of sons were more likely to judge that their sons had talent than parents of daughters. A panoply of objective measures, including math grades in school, performance on standardized tests, teachers' evaluations, and children's expressed interest in math, revealed no differences between the girls and boys. Still, there was a difference in parents' perception of their child's intangible talent. Other studies have shown a similar effect for science.

There's clearly a mismatch between what parents perceive in their kids and what objective measures reveal. But is it possible that the parents are seeing something

that the objective measures are missing? Maybe the boy getting B's in his math class really is a mathematical genius, and his mom or dad has sensed that. To eliminate that possibility, we need to present observers with the very same baby, or child, or Ph.D. candidate, and manipulate their belief about the person's gender. Then we can ask whether their belief influences their perception.

It's hard to do these studies, but there are examples, and I will describe a few of them. A bunch of studies take the following form: you show a group of parents, or college undergraduates, video-clips of babies that they don't know personally. For half of them you give the baby a male name, and for the other half you give the baby a female name. (Male and female babies don't look very different.) The observers watch the baby and then are asked a series of questions: What is the baby doing? What is the baby feeling? How would you rate the baby on a dimension like strong-to-weak, or more intelligent to less intelligent? There are two important findings.

First, when babies do something unambiguous, reports are not affected by the baby's gender. If the baby clearly smiles, everybody says the baby is smiling or happy. Perception of children is not pure hallucination. Second, children often do things that are ambiguous, and parents face questions whose answers aren't easily readable off their child's overt behavior. In those cases, you see some interesting gender labeling effects. For example, in one study a child on a video-clip was playing with a jack-in-the-box. It suddenly popped up, and the child was startled and jumped backward. When people were asked, what's the child feeling, those who were given a female label said, "she's afraid." But the ones given a male label said, "he's angry." Same child, same reaction, different interpretation.

In other studies, children with male names were more likely to be rated as strong, intelligent, and active; those with female names were more likely to be rated as little, soft, and so forth.

I think these perceptions matter. You, as a parent, may be completely committed to treating your male and female children equally. But no sane parents would treat a fearful child the same way they treat an angry child. If knowledge of a child's gender affects adults' perception of that child, then male and female children are going to elicit different reactions from the world, different patterns of encouragement. These perceptions matter, even in parents who are committed to treating sons and daughters alike.

I will give you one last version of a gender-labeling study. This one hits particularly close to home. The subjects in the study were people like Steve and me: professors of psychology, who were sent some vitas to evaluate as applicants for a tenure track position. Two different vitas were used in the study. One was a vita of a walk-on-water candidate, best candidate you've ever seen, you would die to have this person on your faculty. The other vita was a middling, average vita among successful candidates. For half the professors, the name on the vita was male, for the other half the name was female. People were asked a series of questions: What do you think

about this candidate's research productivity? What do you think about his or her teaching experience? And finally, Would you hire this candidate at your university?

For the walk-on-water candidate, there was no effect of gender labeling on these judgments. I think this finding supports Steve's view that we're dealing with little overt discrimination at universities. It's not as if professors see a female name on a vita and think, I don't want her. When the vita's great, everybody says great, let's hire.

What about the average successful vita, though: that is to say, the kind of vita that professors most often must evaluate? In that case, there were differences. The male was rated as having higher research productivity. These psychologists, Steve's and my colleagues, looked at the same number of publications and thought, "good productivity" when the name was male, and "less good productivity" when the name was female. Same thing for teaching experience. The very same list of courses was seen as good teaching experience when the name was male, and less good teaching experience when the name was female. In answer to the question would they hire the candidate, 70% said yes for the male, 45% for the female. If the decision were made by majority rule, the male would get hired and the female would not.

A couple other interesting things came out of this study. The effects were every bit as strong among the female respondents as among the male respondents. Men are not the culprits here. There were effects at the tenure level as well. At the tenure level, professors evaluated a very strong candidate, and almost everyone said this looked like a good case for tenure. But people were invited to express their reservations, and they came up with some very reasonable doubts. For example, "This person looks very strong, but before I agree to give her tenure I would need to know, was this her own work or the work of her adviser?" Now that's a perfectly reasonable question to ask. But what ought to give us pause is that those kinds of reservations were expressed four times more often when the name was female than when the name was male.

So there's a pervasive difference in perceptions, and I think the difference matters. Scientists' perception of the quality of a candidate will influence the likelihood that the candidate will get a fellowship, a job, resources, or a promotion. A pattern of biased evaluation therefore will occur even in people who are absolutely committed to gender equity.

I have little doubt that all my colleagues here at Harvard are committed to the principle that a male candidate and a female candidate of equal qualifications should have equal chance at a job. But we also think that when we compare a more productive scholar to a less productive one, a more experienced teacher to a less experienced one, a more independent investigator to a less independent one, those factors matter as well. These studies say that knowledge of a person's gender will influence our assessment of those factors, and that's going to produce a pattern of discrimination, even in people with the best intentions.

From the moment of birth to the moment of tenure, throughout this great developmental progression, there are unintentional but pervasive and important differences in the ways that males and females are perceived and evaluated.

I have to emphasize that perceptions are not everything. When cases are unambiguous, you don't see these effects. What's more, cognitive development is robust: boys and girls show equal capacities and achievements in educational settings, including in science and mathematics, despite the very different ways in which boys and girls are perceived and evaluated. I think it's really great news that males and females develop along common paths and gain common sets of abilities. The equal performance of males and females, despite their unequal treatment, strongly suggests that mathematical and scientific reasoning has a biological foundation, and this foundation is shared by males and females.

Finally, you do not create someone who feels like a girl or boy simply by perceiving them as male or female. That's the lesson that comes from the studies of people of one sex who are raised as the opposite sex. Biological sex differences are real and important. Sex is not a cultural construction that's imposed on people.

But the question on the table is not, Are there biological sex differences? The question is, Why are there fewer women mathematicians and scientists? The patterns of bias that I described provide four interconnected answers to that question. First, and most obviously, biased perceptions produce discrimination: When a group of equally qualified men and women are evaluated for jobs, more of the men will get those jobs if they are perceived to be more qualified. Second, if people are rational, more men than women will put themselves forward into the academic competition, because men will see that they've got a better chance for success. Academic jobs will be more attractive to men because they face better odds, will get more resources, and so forth.

Third, biased perceptions earlier in life may well deter some female students from even attempting a career in science or mathematics. If your parents feel that you don't have as much natural talent as someone else whose objective abilities are no better than yours, that may discourage you, as Eccles's work shows. Finally, there's likely to be a snowball effect. All of us have an easier time imagining ourselves in careers where there are other people like us. If the first three effects perpetuate a situation where there are few female scientists and mathematicians, young girls will be less likely to see math and science as a possible life.

So by my personal scorecard, these are the major factors. Let me end, though, by asking, could Steve also be partly right? Could biological differences in motives — motivational patterns that evolved in the Pleistocene but that apply to us today — propel more men than women towards careers in mathematics and science?

My feeling is that where we stand now, we cannot evaluate this claim. It may be true, but as long as the forces of discrimination and biased perceptions affect people so pervasively, we'll never know. I think the only way we can find out is to do one more

experiment. We should allow all of the evidence that men and women have equal cognitive capacity, to permeate through society. We should allow people to evaluate children in relation to their actual capacities, rather than one's sense of what their capacities ought to be, given their gender. Then we can see, as those boys and girls grow up, whether different inner voices pull them in different directions. I don't know what the findings of that experiment will be. But I do hope that some future generation of children gets to find out.

Steven Pinker & Elizabeth Spelke: Concluding Discussion

PINKER: Thanks, Liz, for a very stimulating and apposite presentation. A number of comments.

I don't dispute a lot of the points you made, but many have lost sight of the datum that we're here to explain in the first place. Basic abilities like knowing that an object is still there when you put a hankie over it, or knowing that one object can't pass through another, are not the kinds of things that distinguish someone who's capable of being a professor of physics or math from someone who isn't. And in many of the cases in which you correctly said that there is no gender difference in kids, there is no gender difference in adults either — such as the give-a-number task and other core abilities.

Also, a big concern with all of the null effects that you mentioned is statistical power. Bob Rosenthal 20 years ago pointed out that the vast majority of studies that psychologists do are incapable of detecting the kinds of results they seek, which is why it's so important to have meta-analyses and large sample sizes. I question whether all of the null results that you mentioned can really be justified, and whether they are comparable to the studies done on older kids and adults.

One place where I really do disagree with you is in the value of the SAT-M, where the "circle" has amply been broken. This is what people at the College Board are obsessed with. What you are treating as the gold standard is performance in college courses. But the datum we are disputing is not how well boys and girls do in school, or how well men and women do in college, because there we agree there is no male advantage. The phenomenon we really are discussing is performance at the upper levels: getting tenure-track job, getting patents, and so on. And here the analyses have shown that the SAT is not biased against girls. That is, a given increment in SAT score predicts a given increment in the variable of interest to the same extent whether you're male or female.

I think there may be a slight difference in which finding each of us is alluding to in talking about differences in grades. I was not suggesting that girls' better grades come about because they take easier courses; they really do get better grades

holding courses constant. Rather it's the slight underprediction of grades by the SAT that can be explained in part by class choice and in part by conscientiousness.

SPELKE: Well the most recent thing that I've read about this issue is the Gallagher and Kaufman book, Gender Differences in Mathematics, which just came out about a month ago. They report that equating for classes and institutions, and looking just at A students, there's a 21 point SAT math differential; that is to say, for two students getting the same grade of A, the average for the girls on the SAT will have been 21 points lower. That differential is there at every grade level and in all the courses.

The SAT people have discussed it as a problem. One of the discussions reached the conclusion that the SAT is still useful, because although it under-predicts girls' performance in college, girls' grades over-predict their performance in college, and if you use the two together you are okay. In fact, they advised that people never take account of the SAT simply by itself, but consider it in relation to grades. When you spoke earlier about the use of GREs in admitting people to grad school, that's in fact what graduate programs do: We consider both grades and GREs.

Interestingly, though, in all of the public discussion of the relative advantages of men versus women for math and science, over the last two months, people have not used the SAT in conjunction with grades. When talking about relative ability, they've used the SAT by itself. I think that has led to a distorted conversation about this issue.

PINKER: It nonetheless remains true that in the most recent study by Lubinski and Benbow, which showed a fantastic degree of predictive power of the SAT given in 7th grade, there was no difference in predictive power in boys and girls in any of these measures.

But let me return to the datum that is at issue here, namely the differential representation of the sexes in physical sciences, mechanical engineering, and mathematics. The fact that men and women are equal overall in spatial abilities, and overall in mathematical abilities, is irrelevant to this. It may be that the particular subtalents in which women excel make them more likely to go into accounting. But the datum we are discussing is not a gender difference in accounting. The datum we are discussing is a gender difference in the physical sciences, engineering, and mathematics. And I suspect that when you look at a range of professions, the size of the sex discrepancy correlates with how much spatial manipulation (not just any kind of spatial cognition) and how much mathematical reasoning (not just any kind of mathematical ability) each of those jobs requires.

What about parents' expectations? In the 1970s the model for development was, "as the twig is bent, so grows the branch." — that subtle differences in parents' perceptions early in life can have a lasting effect. You nudge the child in a particular direction and you'll see an effect on his trajectory years later. But there is now an enormous amount of research spearheaded by the behavioral genetics revolution suggesting that that is not true. There may be effects of parental expectations and

parental treatment on young children while they're still in the home, but most follow-up studies show that short of outright abuse and neglect, these effects peter out by late adolescence. And studies of adoption and of twins and other sibs reared apart suggest that any effects of the kinds of parenting that are specific to a child simply reflect the preexisting genetic traits of the child, and the additional effect of parenting peters out to nothing.

SPELKE: Can I respond to that? I think one thing is different about the gender case, compared to the early socialization effects for other kinds of categories, different styles of parenting, and so forth. The gender differences that we see reflected in parents' differing perceptions are mirrored by differing perceptions that males and females experience throughout their lives. It's not the case that idiosyncratic pairs of parents treat their kids one way, but then as soon as the children leave that environment, other people treat them differently. Rather, what we have in the case of gender is a pervasive pattern that just keeps getting perpetuated in different people. I'm rather a nativist about cognition, and I am tempted to look at that pattern and wonder, did Darwin give us some innately wrong idea about the genders? Professionals in professional contexts show the same patterns of evaluation that parents show in home contexts, and children face those patterns of evaluation, not just when they're young and at home, but continuing through high school, college, and finally with their colleagues on academic faculties. We're dealing here with a much more pervasive effect than the effects of socialization in the other studies that you've written and talked about.

PINKER: Regarding bias: as I mentioned at the outset, I don't doubt that bias exists. But the idea that the bias started out from some arbitrary coin flip at the dawn of time and that gender differences have been perpetuated ever since by the existence of that bias is extremely unlikely. In so many cases, as Eagly and the Stereotype-Accuracy people point out, the biases are accurate. Also, there's an irony in these discussion of bias. When we test people in the cognitive psychology lab, and we don't call these base rates "gender," we applaud people when they apply them. If people apply the statistics of a group to an individual case, we call it rational Bayesian reasoning, and congratulate ourselves for getting them to overcome the cognitive illusion of base rate neglect. But when people do the same thing in the case of gender, we treat Bayesian reasoning as a cognitive flaw and base-rate neglect as rational! Now I agree that applying base rates for gender in evaluating individual men and women is a moral flaw; I don't think that base rates ought to be applied in judging individuals in most cases of public decision-making. But the fact that the statistics of a gender are applied does not mean that their origin was arbitrary; it could be statistically sound in some cases.

SPELKE: Let me reply to that, because I agree that the origin is not arbitrary, and that the bias is there for an objective reason, but I think you're drawing the wrong conclusion about it. I think the reason there's a bias to think that men have greater natural talent for math and science is that when we look around the world and ask,

who's winning the Nobel Prizes and making the great advances in science, what we see, again and again, is men.

Although Linda Buck received this year's Nobel Prize in physiology or medicine, for the most part it's overwhelmingly men who are reaching the upper levels of math and science. It's natural to look at that and think, there must be some reason, some inner difference between men and women, which produces this enormous disparity. And I quite agree with you that good statistical reasoning should lead you to think, the next student who comes along, if male, is more likely to join that group of Nobel Prize winners.

What I would like to suggest is that we have good reasons to resist this kind of conclusion, and the reasons aren't only moral. Let me just use an analogy, and replay this debate over the biological bases of mathematics and science talent 150 years ago.

Let's consider who the 19th century mathematicians and scientists were. They were overwhelmingly male, just as they are today, but also overwhelmingly European, not Asian. You won't see a Chinese face or an Indian face in 19th century science. It would have been tempting to apply this same pattern of statistical reasoning and say, there must be something about European genes that give rise to greater mathematical talent than Asian genes do. If we go back still further, and play this debate in the Renaissance, I think we would be tempted to conclude that Catholic genes make for better science than Jewish genes, because all those Renaissance scientists were Catholic. If you look at those cases, you see what's wrong with this argument.

What's wrong with the argument is not that biology is irrelevant. If Galileo had been switched at birth with some baby from the Pisan ghetto, the baby raised by Galileo's parents would not likely have ended up teaching us that the language of physics is mathematics. I think that Galileo's genes had something to do with his achievement, but so did Galileo's cultural and social environment: his nurturing. Genius requires huge amounts of both. If, in that baby switch, Galileo had found himself growing up in the Pisan ghetto, I bet he wouldn't have ended up being the example in this discussion today either. So yes, there are reasons for this statistical bias. But I think we want to step back and ask, why is it that almost all Nobel Prize winners are men today? The answer to that question may be the same reason why all the great scientists in Florence were Christian.

PINKER: I think you could take the same phenomenon and come to the opposite conclusion! Say there were really was such a self-reinforcing, self-perpetuating dynamic: a difference originates for reasons that might be arbitrary; people perceive the difference; they perpetuate it by their expectations. Just as bad, you say, is the fact that people don't go into fields in which they don't find enough people like themselves. If so, the dynamic you would expect is that the representation of different genders or ethnic groups should migrate to the extremes. That is, there is a positive feedback loop where if you're in the minority, it will discourage people like

you from entering the field, which will mean that there'll be even fewer people in the field, and so on. On either side of this threshold you should get a drift of the percentages in opposite directions.

Now, there is an alternative model. At many points in history, arbitrary barriers against the entry of genders and races and ethnic groups to various professions were removed. And as soon as the barrier was removed, far from the statistical underrepresentation perpetuating or exaggerating itself, as you predict, the floodgates open, and the formerly underrepresented people reaches some natural level. It's the Jackie Robinson effect in baseball. In the case of gender and science, remember what our datum is. It's not that women are under-represented in professions in general or in the sciences in general: in many professions women are perfectly well represented, such as being a veterinarian, in which the majority of recent graduates are women by a long shot. If you go back fifty years or a hundred years, there would have been virtually no veterinarians who were women. That underrepresentation did not perpetuate itself via the positive feedback loop that you allude to.

SPELKE: I'm glad you brought up the case of the basketball and baseball players. I think it's interesting to ask, what distinguishes these cases, where you remove the overt discrimination and within a very short period of time the differential disappears, from other cases, where you remove the overt discrimination and the covert discrimination continues? In the athletic cases where discrimination disappears quickly, there are clear, objective measures of success. Whatever people think about the capacities of a black player, if he is hitting the ball out of the park, he is going to get credit for a home run. That is not the case in science.

In science, the judgments are subjective, every step of the way. Who's really talented? Who deserves bigger lab space? Who should get the next fellowship? Who should get promoted to tenure? These decisions are not based on clear and objective criteria. These are the cases where you see discrimination persisting. You see it in academia. You see it in Claudia Goldin's studies of orchestra auditions, which also involve subtle judgments: Who's the more emotive, sensitive player? If you know that the players are male or female, you're going pick mostly men, but if the players are behind a screen, you'll start picking more women.

PINKER: But that makes the wrong prediction: the harder the science, the greater the participation of women! We find exactly the opposite: it's the most subjective fields within academia — the social sciences, the humanities, the helping professions — that have the greatest representation of women. This follows exactly from the choices that women express in what gives them satisfaction in life. But it goes in the opposite direction to the prediction you made about the role of objective criteria in bringing about gender equity. Surely it's physics, and not, say, sociology, that has the more objective criteria for success.

SPELKE: Let me just say one thing, because I didn't say much in the talk at all, about this issue of motives, and biological differences in motives. That's been a less

controversial issue, but I think it's an important one, and most of your examples were concerned with it. I think it's a really interesting possibility that the forces that were active in our evolutionary past have led men and women to evolve somewhat differing concerns. But to jump from that possibility into the present, and draw conclusions about what people's motives will be for pursuing one or another career, is way too big a stretch.

As we both agree, the kinds of careers people pursue now, the kinds of choices they make, are radically different from anything that anybody faced back in the Pleistocene. It is anything but clear how motives that evolved then translate into a modern context. Let me just give one example of this. You've suggested, as a hypothesis, that because of sexual selection and also parental investment issues, men are selected to be more competitive, and women are selected to be more nurturant. Suppose that hypothesis is true. If we want to use it to make predictions about desires for careers in math and science, we're going to have to answer a question that I think is wide open right now. What makes for better motives in a scientist?

What kind of motives are more likely to lead to good science: Competitive motives, like the motive J. D. Watson described in The Double Helix, to get the structure of DNA before Linus Pauling did? Or nurturant motives of the kind that Doug Melton has described recently to explain why he's going into stem cell research: to find a cure for juvenile diabetes, which his children suffer from? I think it's anything but clear how motives from our past translate into modern contexts. We would need to do the experiment, getting rid of discrimination and social pressures, in order to find out.

(Links et al., in discussion below)

To: David Kroenlein:

With daughters and grand-daughters you might be interested. I like academic jousting seeking truth as we can know it, but often it's like lawyers, taking fact or factoid here and there to advocate a position. The Pinker/Spelke thing via the link is a good read. Good friend Ed has picked up on the Summers thing I sent and done well with it.

In middle of this exchange I had also emailed this to him:

Hopkins thing I thought came out years ago but maybe not. Reaction has been similar to those Princeton economists suggesting elite school education adds little in earning power over the years for the same reason: intelligentsia who are in the system don't like implications re their own merit and/or don't know what to do about it anyway. E.g., if tenure process is flawed, there are others who deserved to be there.....and maybe i did not!

The research is almost always subject to differing conclusions or challenge Given how the tested hypothesis is phrased. Leftists liked sexual openness and didn't like violence decades ago. Conclusion in multiple studies was that lots of TV watching by young males increased their tolerance and desire for violence yet had no effect on their sexual experimentations, frequency etc. (this, of course, was before the fad against pornography inducing violence against women.). That's likely nonsense on the face of it.

I don't think this stuff convinces me that the tighter distribution theory is disproved. But for safety, I agree the point is to be born a WASP male.

Social "science" research can prove nothing but can only say things are consistent or inconsistent with an hypothesis, Milton Friedman and many other Chicago economists notwithstanding. We still have to act. That's the rub.

From: "E. Eugene CARTER" <eecarter1@gmail.com>

I really enjoyed plowing thru the pinker/Spelke discussion. First rate minds are seen driving to support their conclusions like good lawyers, overlaid with civility and quest for truth., unlike litigators. Spelke, like a Chicago economist, is dirtier, citing some null results as comparably significant, pushing grades as measures (which you buy and Pinker and I don't), etc. She seems more often to overstate conclusions than he does. I agree it's complex, and I don't know the studies, and know each participant selects them. We can get a print out and go over at lunch sometime. I just find pinker less overstating what he finds and can argue than she does. That makes me think she is more the advocate than the analyst at this point. I am more convinced a lot is the fat tails.

As I noted and they allude to, a lot is in the nature of social science conclusions, and the softness in anything. He seems a bit more circumspect whereas she grabs conclusions with less support, or extends beyond what a study shows. Everybody has to simplify and generalize in a talk of course. The meta studies are there and do support lots of toy pref differences, task prefs, etc. She just loves to fall back on discrimination and socialization every chance she gets, which is tedious and not as helpful. "There's no difference and any thing you find is perception bias or discrimination." Great. It's there. Given a fixation I can use it to explain away most studies, etc. That's not what she's doing, but you see it with Chicago economists and MIT finance types. It's just not robust science by its nature and cannot be.

On Jun 9, 2010, at 12:43 PM, "Ed Lang" <ed@lang.net> wrote:

Gene,

Spelke finds Summers/Pinker "tighter distribution" and "right-tail" hypotheses flawed by Sat-M circularity, but you will see that when you look at her presentation.

What I found especially riveting came later, at 1AM, as I watched Nancy Hopkins give a lecture:

http://www.princeton.edu/WebMedia/flash/lectures/20090331_publect_hopkins.shtml

Hopkins described how it took her so many years to understand what was happening. How she thought she was the problem. How she thought women not being willing to put in the time, was the problem. How she thought gender bias was nonsense.

When she collected data about women at MIT, the data led, over time, to a conclusion she didn't want to accept. She was happier before, when she thought she and time were the problem. It was a 40-year process to understand the problem and it's implications.

She came to understand that unintentional, implicit gender bias was the problem -- that both men and women think and judge the performance of men to be slightly better than women. Empirical evidence from experimental psychology shows it to be true. It's pervasive, it's natural, it may start happening in children younger than one year old.

I felt sadness by the end of the lecture.

-Ed

On Tue, Jun 8, 2010 at 10:42 PM, Ed Lang <ed@lang.net> wrote:

Gene,

Your email prompted me to look into the issue of gender in science, and I found an interesting 2005 Pinker-Spelke presentation on the subject.

http://www.edge.org/3rd_culture/debate05/debate05_index.html

I thought Spelke quite persuasive in showing:

* unintentional perceptual gender bias against women when evaluating good (but not outstanding) job candidates.

* no empirical evidence supports greater male intrinsic aptitude for mathematics or science.

leading to a conclusion:

* gender inequality in math and science is mostly a consequence of discrimination (from perceptual bias) and social forces.

Our evaluations and judgments are all based on perception, and empirical data shows perception is biased. A big problem if we deny it, because it's true. Perceptual bias is unavoidable, for everyone. If we value fairness and equality, how do we achieve it?

Curious what you think.

- Ed

June 7, 2010 NY TIMES

Daring to Discuss Women in Science

By JOHN TIERNEY

The House of Representatives has passed what I like to think of as Larry's Law. The official title of this legislation is "Fulfilling the potential of women in academic science and engineering," but nothing did more to empower its advocates than the controversy over a speech by Lawrence H. Summers when he was president of Harvard.

This proposed law, if passed by the Senate, would require the White House science adviser to oversee regular "workshops to enhance gender equity." At the workshops, to be attended by researchers who receive federal money and by the heads of science and engineering departments at universities, participants would be given before-and-after "attitudinal surveys" and would take part in "interactive discussions or other activities that increase the awareness of the existence of gender bias."

I'm all in favor of women fulfilling their potential in science, but I feel compelled, at the risk of being shipped off to one of these workshops, to ask a couple of questions:

1) Would it be safe during the "interactive discussions" for someone to mention the new evidence supporting Dr. Summers's controversial hypothesis about differences in the sexes' aptitude for math and science?

2) How could these workshops reconcile the "existence of gender bias" with careful studies that show that female scientists fare as well as, if not better than, their male counterparts in receiving academic promotions and research grants?

Each of these questions is complicated enough to warrant a column, so I'll take them one at a time, starting this week with the issue of sex differences.

When Dr. Summers raised the issue to fellow economists and other researchers at a conference in 2005, his hypothesis was caricatured in the press as a revival of the old notion that "girls can't do math." But Dr. Summers said no such thing. He acknowledged that there were many talented female scientists and discussed ways to eliminate the social barriers they faced.

Yet even if all these social factors were eliminated, he hypothesized, the science faculty composition at an elite school like Harvard might still be skewed by a biological factor: the greater variability observed among men in intelligence test scores and various traits. Men and women might, on average, have equal

mathematical ability, but there could still be disproportionately more men with very low or very high scores.

These extremes often don't matter much because relatively few people are involved, leaving the bulk of men and women clustered around the middle. But a tenured physicist at a leading university, Dr. Summers suggested, might well need skills and traits found in only one person in 10,000: the top 0.01 percent of the population, a tiny group that would presumably include more men because it's at the extreme right tail of the distribution curve.

"I would like nothing better than to be proved wrong," Dr. Summers told the economists, expressing the hope that gender imbalances could be rectified simply by eliminating social barriers. But he added, "My guess is that there are some very deep forces here that are going to be with us for a long time."

Dr. Summers was pilloried for even suggesting the idea, and the critics took up his challenge to refute the hypothesis. Some have claimed he was proved wrong by recent reports of girls closing the gender gap on math scores in the United States and other countries. But even if those reports (which have been disputed) are accurate, they involve closing the gap only for average math scores — not for the extreme scores that Dr. Summers was discussing.

Some scientists and advocates for gender equity have argued that the remaining gender gap in extreme scores is rapidly shrinking and will disappear. It was called "largely an artifact of changeable sociocultural factors" last year by two researchers at the University of Wisconsin, Janet S. Hyde and Janet E. Mertz. They noted evidence of the gap narrowing and concluded, "Thus, there is every reason to believe that it will continue to narrow in the future."

But some of the evidence for the disappearing gender gap involved standardized tests that aren't sufficiently difficult to make fine distinctions among the brighter students. These tests, like the annual ones required in American public schools, are limited by what's called the ceiling effect: If you're measuring people in a room with a six-foot ceiling, you can't distinguish among the ones taller than six feet.

Now a team of psychologists at Duke University has looked at the results of tests with more headroom. In an article in a forthcoming issue of the journal Intelligence, they analyze the test scores of students in the United States who took college admissions tests while they were still in the seventh grade. As part of an annual talent search since 1981, the SAT and ACT tests have been given to more than 1.6 million gifted seventh graders, with roughly equal numbers of boys and girls participating.

The Duke researchers — Jonathan Wai, Megan Cacchio, Martha Putallaz and Matthew C. Makel — focused on the extreme right tail of the distribution curve: people ranking in the top 0.01 percent of the general population, which for a seventh grader means scoring above 700 on the SAT math test. In the early 1980s, there were 13 boys for every girl in that group, but by 1991 the gender gap had

narrowed to four to one, presumably because of sociocultural factors like encouragement and instruction in math offered to girls.

Since then, however, the math gender gap hasn't narrowed, despite the continuing programs to encourage girls. The Duke researchers report that there are still four boys for every girl at the extreme right tail of the scores for the SAT math test. The boy-girl ratio has also remained fairly constant, at about three to one, at the right tail of the ACT tests of both math and science reasoning. Among the 19 students who got a perfect score on the ACT science test in the past two decades, 18 were boys.

Meanwhile, the seventh-grade girls outnumbered the boys at the right tail of tests measuring verbal reasoning and writing ability. The Duke researchers report in Intelligence, "Our data clearly show that there are sex differences in cognitive abilities in the extreme right tail, with some favoring males and some favoring females."

The researchers say it's impossible to predict how long these math and science gender gaps will last. But given the gaps' stability for two decades, the researchers conclude, "Thus, sex differences in abilities in the extreme right tail should not be dismissed as no longer part of the explanation for the dearth of women in math-intensive fields of science."

Other studies have shown that these differences in extreme test scores correlate with later achievements in science and academia. Even when you consider only members of an elite group like the top percentile of the seventh graders on the SAT math test, someone at the 99.9 level is more likely than someone at the 99.1 level to get a doctorate in science or to win tenure at a top university.

Of course, a high score on a test is hardly the only factor important for a successful career in science, and no one claims that the right-tail disparity is the sole reason for the relatively low number of female professors in math-oriented sciences. There are other potentially more important explanations, both biological and cultural, including possible social bias against women.

But before we accept Congress's proclamation of bias, before we start re-educating scientists at workshops, it's worth taking a hard look at the evidence of bias against female scientists. That will be the subject of another column.