



[PRINT WINDOW](#) [CLOSE WINDOW](#)

BELL CURVE

The Inequality Taboo

It's time to start talking about differences between groups of people.

BY CHARLES MURRAY

Wednesday, October 12, 2005 12:01 a.m.

When the late Richard Herrnstein and I published "The Bell Curve" 11 years ago, the furor over its discussion of ethnic differences in IQ was so intense that most people who have not read the book still think it was about race. Since then, I have deliberately not published anything about group differences in IQ, mostly to give the real topic of "The Bell Curve"--the role of intelligence in reshaping America's class structure--a chance to surface.

The Lawrence Summers affair last January made me rethink my silence. The president of Harvard University offered a few mild, speculative, off-the-record remarks about innate differences between men and women in their aptitude for high-level science and mathematics, and was treated by Harvard's faculty as if he were a crank. The typical news story portrayed the idea of innate sex differences as a renegade position that reputable scholars rejected.

It was depressingly familiar. In the autumn of 1994, I had watched with dismay as "The Bell Curve" 's scientifically unremarkable statements about black IQ were successfully labeled as racist pseudoscience. At the opening of 2005, I watched as some scientifically unremarkable statements about male-female differences were successfully labeled as sexist pseudoscience.



The Orwellian disinformation about innate group differences is not wholly the media's fault. Many academics who are familiar with the state of knowledge are afraid to go on the record. Talking publicly can dry up research funding for senior professors and can cost assistant professors their jobs. But while the public's misconception is

understandable, it is also getting in the way of clear thinking about American social policy.

Good social policy can be based on premises that have nothing to do with scientific truth. The premise that is supposed to undergird all of our social policy, the founders' assertion of an unalienable right to liberty, is not a falsifiable hypothesis. But specific policies based on premises that conflict with scientific truths about human beings tend not to work. Often they do harm.

One such premise is that the distribution of innate abilities and propensities is the same across different groups. The statistical tests for uncovering job discrimination assume that men are not innately different from women, blacks from whites, older people from younger people, homosexuals from heterosexuals, Latinos from Anglos, in ways that can legitimately affect employment decisions. Title IX of the Educational Amendments of 1972 assumes that women are no different from men in their attraction to sports. Affirmative action in all its forms assumes there are no innate differences between any of the groups it seeks to help and everyone else. The assumption of no innate differences among groups suffuses American social policy. That assumption is wrong.

When the outcomes that these policies are supposed to produce fail to occur, with one group falling short, the fault for the discrepancy has been assigned to society. It continues to be assumed that better programs, better regulations or the right court decisions can make the differences go away. That assumption is also wrong.

Hence this essay. Most of the following discussion describes reasons for believing that some group differences are intractable. I shift from "innate" to "intractable" to acknowledge how complex is the interaction of genes, their expression in behavior, and the environment. "Intractable" means that, whatever the precise partitioning of causation may be (we seldom know), policy interventions can only tweak the difference at the margins.

I will focus on two sorts of differences: between men and women and between blacks and whites. Here are three crucial points to keep in mind as we go along:

1. The differences I discuss involve means and distributions. In all cases, the variation within groups is greater than the variation between groups. On psychological and cognitive dimensions, some members of both sexes and all races fall everywhere along the range. One implication of this is that genius does not come in one color or sex, and neither does any other human ability. Another is that a few minutes of conversation with individuals you meet will tell you much more about them than their group membership does.

2. Covering both sex differences and race differences in a single nontechnical article, I have had to leave out much. I urge that readers with questions consult the fully annotated version of this essay, which includes extensive supplementary material; it is

available [here](#) at Commentary's Web site.

3. The concepts of "inferiority" and "superiority" are inappropriate to group comparisons. On most specific human attributes, it is possible to specify a continuum running from "low" to "high," but the results cannot be combined into a score running from "bad" to "good." What is the best score on a continuum measuring aggressiveness? What is the relative importance of verbal skills versus, say, compassion? Of spatial skills versus industriousness? The aggregate excellences and shortcomings of human groups do not lend themselves to simple comparisons. That is why the members of just about every group can so easily conclude that they are God's chosen people. All of us use the weighting system that favors our group's strengths.



The technical literature documenting sex differences and their biological basis grew surreptitiously during feminism's heyday in the 1970s and 1980s. By the 1990s, it had become so extensive that the bibliography in David Geary's pioneering "Male, Female" (1998) ran to 53 pages. Currently, the best short account of the state of knowledge is Steven Pinker's chapter on sex in "The Blank Slate" (2002).

Rather than present a telegraphic list of all the differences that I think have been established, I will focus on the narrower question at the heart of the Summers controversy: As groups, do men and women differ innately in characteristics that produce achievement at the highest levels of accomplishment? I will limit my comments to the arts and sciences.

Since we live in an age when students are likely to hear more about Marie Curie than about Albert Einstein, it is worth beginning with a statement of historical fact: Women have played a proportionally tiny part in the history of the arts and sciences. Others have found similar proportions. Even in the 20th century, women got only 2% of the Nobel Prizes in the sciences--a proportion constant for both halves of the century--and 10% of the prizes in literature. The Fields Medal, the most prestigious award in mathematics, has been given to 44 people since it originated in 1936. All have been men.

The historical reality of male dominance of the greatest achievements in science and the arts is not open to argument. The question is whether the social and legal exclusion of women is a sufficient explanation for this situation, or whether sex-specific characteristics are also at work.

Mathematics offers an entry point for thinking about the answer. Through high school, girls earn better grades in math than boys, but boys usually do better on standardized tests. The difference in means is modest, but the male advantage increases as the focus shifts from means to extremes. In a large sample of mathematically gifted youths, for example, seven times as many males as females scored in the top percentile of the SAT mathematics test. We do not have good test data on the male-female ratio at the top

one-hundredth or top one-thousandth of a percentile, where first-rate mathematicians are most likely to be found, but collateral evidence suggests that the male advantage there continues to increase, perhaps exponentially.

Evolutionary biologists have some theories that feed into an explanation for the disparity. In primitive societies, men did the hunting, which often took them far from home. Males with the ability to recognize landscapes from different orientations and thereby find their way back had a survival advantage. Men who could process trajectories in three dimensions--the trajectory, say, of a spear thrown at an edible mammal--also had a survival advantage. Women did the gathering. Those who could distinguish among complex arrays of vegetation, remembering which were the poisonous plants and which the nourishing ones, also had a survival advantage. Thus the logic for explaining why men should have developed elevated three-dimensional visuospatial skills and women an elevated ability to remember objects and their relative locations--differences that show up in specialized tests today.



Perhaps this is a just-so story. Why not instead attribute the results of these tests to socialization? Enter the neuroscientists. It has been known for years that even after adjusting for body size, men have larger brains than women. Yet most psychometricians conclude that men and women have the same mean IQ (although debate on this issue is growing). One hypothesis for explaining this paradox is that three-dimensional processing absorbs the extra male capacity. In the past few years, magnetic-resonance imaging has refined the evidence for this hypothesis, revealing that parts of the brain's parietal cortex associated with space perception are proportionally bigger in men than in women.

What does space perception have to do with scores on math tests? Enter the psychometricians, who demonstrate that when visuospatial ability is taken into account, the sex difference in SAT math scores shrinks substantially.

Why should the difference be so much greater at the extremes than at the mean? Part of the answer is that men consistently exhibit higher variance than women on all sorts of characteristics, including visuospatial abilities, meaning that there are proportionally more men than women at both ends of the bell curve. Another part of the answer is that someone with a high verbal IQ can easily master the basic algebra, geometry and calculus that make up most of the items in an ordinary math test. Elevated visuospatial skills are most useful for the most difficult items. If males have an advantage in answering those comparatively few really hard items, the increasing disparity at the extremes becomes explicable.

Seen from one perspective, this pattern demonstrates what should be obvious: there is nothing inherent in being a woman that precludes high math ability. But there remains a distributional difference in male and female characteristics that leads to a larger number of men with high visuospatial skills. The difference has an evolutionary rationale, a physiological basis and a direct correlation with math scores.

Now put all this alongside the historical data on accomplishment in the arts and sciences. In test scores, the male advantage is most pronounced in the most abstract items. Historically, too, it is most pronounced in the most abstract domains of accomplishment.

In the humanities, the most abstract field is philosophy--and no woman has been a significant original thinker in any of the world's great philosophical traditions. In the sciences, the most abstract field is mathematics, where the number of great female mathematicians is approximately two (Emmy Noether definitely, Sonya Kovalevskaya maybe). In the other hard sciences, the contributions of great women have usually been empirical rather than theoretical, with leading cases in point being Henrietta Leavitt, Dorothy Hodgkin, Lise Meitner, Irène Joliot-Curie and Marie Curie herself.

In the arts, literature is the least abstract and by far the most rooted in human interaction; visual art incorporates a greater admixture of the abstract; musical composition is the most abstract of all the arts, using neither words nor images. The role of women has varied accordingly. Women have been represented among great writers virtually from the beginning of literature, in East Asia and South Asia as well as in the West. Women have produced a smaller number of important visual artists, and none clearly in the first rank. No female composer is even close to the first rank. Social restrictions undoubtedly damped down women's contributions in all of the arts, but the pattern of accomplishment that did break through is strikingly consistent with what we know about the respective strengths of male and female cognitive repertoires.



Women have their own cognitive advantages over men, many of them involving verbal fluency and interpersonal skills. If this were a comprehensive survey, detailing those advantages would take up as much space as I have devoted to a particular male advantage. But, sticking with my restricted topic, I will move to another aspect of male-female differences that bears on accomplishment at the highest levels of the arts and sciences: motherhood.

Regarding women, men and babies, the technical literature is as unambiguous as everyday experience would lead one to suppose. As a rule, the experience of parenthood is more profoundly life-altering for women than for men. Nor is there anything unique about humans in this regard. Mammalian reproduction generally involves much higher levels of maternal than paternal investment in the raising of children. Among humans, extensive empirical study has demonstrated that women are more attracted to children than are men, respond to them more intensely on an emotional level, and get more and different kinds of satisfactions from nurturing them. Many of these behavioral differences have been linked with biochemical differences between men and women.

Thus, for reasons embedded in the biochemistry and neurophysiology of being female, many women with the cognitive skills for achievement at the highest level also have

something else they want to do in life: have a baby. In the arts and sciences, 40 is the mean age at which peak accomplishment occurs, preceded by years of intense effort mastering the discipline in question. These are precisely the years during which most women must bear children if they are to bear them at all.

Among women who have become mothers, the possibilities for high-level accomplishment in the arts and sciences shrink because, for innate reasons, the distractions of parenthood are greater. To put it in a way that most readers with children will recognize, a father can go to work and forget about his children for the whole day. Hardly any mother can do this, no matter how good her day-care arrangement or full-time nanny may be. My point is not that women must choose between a career and children, but that accomplishment at the extremes commonly comes from a single-minded focus that leaves no room for anything but the task at hand. We should not be surprised or dismayed to find that motherhood reduces the proportion of highly talented young women who are willing to make that trade-off.

Some numbers can be put to this observation through a study of nearly 2,000 men and women who were identified as extraordinarily talented in math at age 13 and were followed up 20 years later. The women in the sample came of age in the 1970s and early 1980s, when women were actively socialized to resist gender stereotypes. In many ways, these talented women did resist. By their early 30s, both the men and women had become exceptional achievers, receiving advanced degrees in roughly equal proportions. Only about 15% of the women were full-time housewives. Among the women, those who did and those who did not have children were equally satisfied with their careers.

And yet. The women with careers were 4.5 times as likely as men to say they preferred to work less than 40 hours a week. The men placed greater importance on "being successful in my line of work" and "inventing or creating something that will have an impact," while the women found greater value in "having strong friendships," "living close to parents and relatives" and "having a meaningful spiritual life." As the authors concluded, "these men and women appear to have constructed satisfying and meaningful lives that took somewhat different forms." The different forms, which directly influence the likelihood that men will dominate at the extreme levels of achievement, are consistent with a constellation of differences between men and women that have biological roots.

I have omitted perhaps the most obvious reason why men and women differ at the highest levels of accomplishment: Men take more risks, are more competitive and are more aggressive than women. The word *testosterone* may come to mind, and appropriately. Much technical literature documents the hormonal basis of personality differences that bear on sex differences in extreme and venturesome effort, and hence in extremes of accomplishment--and that bear as well on the male propensity to produce an overwhelming proportion of the world's crime and approximately 100% of its wars.

But this is just one more of the ways in which science is demonstrating that men and women are really and truly different, a fact so obvious that only intellectuals could ever

have thought otherwise.



Turning to race, we must begin with the fraught question of whether it even exists, or whether it is instead a social construct. The Harvard geneticist Richard Lewontin originated the idea of race as a social construct in 1972, arguing that the genetic differences across races were so trivial that no scientist working exclusively with genetic data would sort people into blacks, whites or Asians. In his words, "racial classification is now seen to be of virtually no genetic or taxonomic significance."

Mr. Lewontin's position, which quickly became a tenet of political correctness, carried with it a potential means of being falsified. If he was correct, then a statistical analysis of genetic markers would not produce clusters corresponding to common racial labels.

In the past few years, that test has become feasible, and now we know that Mr. Lewontin was wrong. Several analyses have confirmed the genetic reality of group identities going under the label of race or ethnicity. In the most recent, published this year, all but five of the 3,636 subjects fell into the cluster of genetic markers corresponding to their self-identified ethnic group. When a statistical procedure, blind to physical characteristics and working exclusively with genetic information, classifies 99.9% of the individuals in a large sample in the same way they classify themselves, it is hard to argue that race is imaginary.

Homo sapiens actually falls into many more interesting groups than the bulky ones known as "races." As new findings appear almost weekly, it seems increasingly likely that we are just at the beginning of a process that will identify all sorts of genetic differences among groups, whether the groups being compared are Nigerian blacks and Kenyan blacks, lawyers and engineers, or Episcopalians and Baptists. At the moment, the differences that are obviously genetic involve diseases (Ashkenazi Jews and Tay-Sachs disease, black Africans and sickle-cell anemia, Swedes and hemochromatosis). As time goes on, we may yet come to understand better why, say, Italians are more vivacious than Scots.



Out of all the interesting and intractable differences that may eventually be identified, one in particular remains a hot button like no other: the IQ difference between blacks and whites. What is the present state of our knowledge about it?

There is no technical dispute on some of the core issues. In the aftermath of "The Bell Curve," the American Psychological Association established a task force on intelligence whose report was published in early 1996. The task force reached the same conclusions as "The Bell Curve" on the size and meaningfulness of the black-white difference. Historically, it has been about one standard deviation in magnitude among subjects who

have reached adolescence; cultural bias in IQ tests does not explain the difference; and the tests are about equally predictive of educational, social and economic outcomes for blacks and whites. However controversial such assertions may still be in the eyes of the mainstream media, they are not controversial within the scientific community. (The standard deviation is a statistic that, to simplify slightly, expresses the average difference of all the scores from the mean. Given a normal distribution--a bell curve--someone who is one standard deviation above the mean is at the 84th percentile. Two standard deviations above the mean puts that person at the 98th percentile. IQ tests are normed to have a mean of 100 and a standard deviation of 15.)

The most important change in the state of knowledge since the mid-1990s lies in our increased understanding of what has happened to the size of the black-white difference over time. Both the task force and "The Bell Curve" concluded that some narrowing had occurred since the early 1970s. With the advantage of an additional decade of data, we are now able to be more precise: (1) The black-white difference in scores on educational achievement tests has narrowed significantly. (2) The black-white convergence in scores on the most highly "g-loaded" tests--the tests that are the best measures of cognitive ability--has been smaller, and may be unchanged, since the first tests were administered 90 years ago.

With regard to the difference in educational achievement, the narrowing of scores on major tests occurred in the 1970s and '80s. In the case of the SAT, the gaps in the verbal and math tests as of 1972 were 1.24 and 1.26 standard deviations respectively. By 1991, when the gaps were smallest (they have risen slightly since then), those numbers had dropped by 0.37 and 0.35 standard deviation.

The National Assessment of Educational Progress, which is not limited to college-bound students, is preferable to the SAT for estimating nationally representative trends, but the story it tells is similar. Among students ages 9, 13 and 17, the black-white differences in math as of the first NAEP test in 1973 were 1.03, 1.29 and 1.24 standard deviations respectively. For 9-year-olds, the difference hit its all-time low of 0.73 standard deviation in 2004, a drop of 0.30. But almost all of that convergence had been reached by 1986, when the gap was 0.78 standard deviation. For 13-year-olds, the gap dropped by 0.45 standard deviation, reaching its low in 1986. For 17-year-olds, the gap dropped by 0.52 standard deviation, reaching its low in 1990.

In the reading test, the comparable gaps for ages 9, 13 and 17 as of the first NAEP test in 1971 were 1.12, 1.17 and 1.25 standard deviations. Those gaps had shrunk by 0.38, 0.62 and 0.68 standard deviation respectively at their lowest points in 1988. They have since remained effectively unchanged.

An analysis by Larry Hedges and Amy Nowell uses a third set of data, examining the trends for high-school seniors by comparing six large data bases from different time periods from 1965 to 1992. The black-white difference on a combined measure of math, vocabulary and reading fell from 1.18 to 0.82 standard deviation in that time, a reduction of 0.36.



So black and white academic achievement converged significantly in the 1970s and 1980s, typically by more than a third of a standard deviation, and since then has stayed about the same. What about convergence in tests explicitly designed to measure IQ rather than academic achievement? The ambiguities in the data leave two defensible positions. The first is that the IQ difference is about one standard deviation, effectively unchanged since the first black-white comparisons 90 years ago. The second is that harbingers of a narrowing difference are starting to emerge. I cannot settle the argument here, but I can convey some sense of the uncertainty.

The case for an unchanged black-white IQ difference is straightforward. If you take all the black-white differences on IQ tests from the first ones in World War I up to the present, there is no statistically significant downward trend. Of course the results vary, because tests vary in the precision with which they measure the general mental factor (*g*) and samples vary in their size and representativeness. But results continue to center on a black-white difference of about 1.0 to 1.1 standard deviations through the most recent data.

The case for a reduction has two important recent results to work with. The first is from the 1997 renorming of the Armed Forces Qualification Test, which showed a black-white difference of 0.97 standard deviation. Since the typical difference on paper-and-pencil IQ tests like the AFQT has been about 1.10 standard deviations, the 1997 results represent noticeable improvement.

The second positive result comes from the 2003 standardization sample for the Wechsler Intelligence Scale for Children, which showed a difference of 0.78 standard deviation, as against the 1.0 difference that has been typical for individually administered IQ tests.

One cannot draw strong conclusions from two data points. Those who interpret them as part of an unchanging overall pattern can cite another recent result, from the 2001 standardization of the Woodcock-Johnson intelligence test. In line with the conventional gap, it showed an overall black-white difference of 1.05 standard deviations and, for youths age 6 to 18, a difference of 0.99 standard deviation.

There is more to be said on both sides of this issue, but nothing conclusive. Until new data become available, you may take your choice. If you are a pessimist, the gap has been unchanged at about one standard deviation. If you are an optimist, the IQ gap has decreased by a few points, but it is still close to one standard deviation. The clear and substantial convergence that occurred in academic tests has at best been but dimly reflected in IQ scores, and at worst not reflected at all.



Whether we are talking about academic achievement or about IQ, are the causes of the black-white difference environmental or genetic? Everyone agrees that environment plays a part. The controversy is about whether biology is also involved.

It has been known for many years that the obvious environmental factors such as income, parental occupation and schools explain only part of the absolute black-white difference and none of the relative difference. Black and white students from affluent neighborhoods are separated by as large a proportional gap as are blacks and whites from poor neighborhoods. Thus the most interesting recent studies of environmental causes have worked with cultural explanations instead of socioeconomic status. (I put aside here the explanation that has received the most publicity in recent years, the phenomenon labeled "stereotype threat." Its discoverers, Claude Steele and Joshua Aronson, demonstrated experimentally that test performance by academically talented blacks was worse when a test was called an IQ test than when it was innocuously described as a research tool. Press reports erroneously interpreted this as meaning that stereotype threat explained away the black-white difference. In reality, Messrs. Steele and Aronson showed only that it increases the usual black-white difference; if one eliminates stereotype threat, the usual difference remains.)

One example of a cultural explanation is "Black American Students in an Affluent Suburb: A Study of Academic Disengagement" (2003) by the Berkeley anthropologist John Ogbu, who went to Shaker Heights, Ohio, to explore why black students in an affluent suburb should lag behind their white peers. Another is "Black Rednecks and White Liberals" (2005) by Thomas Sowell, who makes the case that what we think of as the dysfunctional aspects of urban black culture are a legacy not of slavery but of Southern and rural white "cracker" culture. Both Mr. Ogbu and Sowell Mr. describe ingrained parental behaviors and student attitudes that must impede black academic performance. These cultural influences often cut across social classes.

From a theoretical standpoint, the cultural explanations offer fresh ways of looking at the black-white difference at a time when the standard socioeconomic explanations have reached a dead end. From a practical standpoint, however, the cultural explanations point to a cause of the black-white difference that is as impervious to manipulation by social policy as causes rooted in biology. If there is to be a rapid improvement, some form of mass movement with powerful behavioral consequences would have to occur within the black community. Absent that, the best we can hope for is gradual cultural change that is likely to be measured in decades.



This brings us to the state of knowledge about genetic explanations. "There is not much direct evidence on this point," said the American Psychological Association's task force dismissively, "but what little there is fails to support the genetic hypothesis." Actually, there is no direct evidence at all, just a wide variety of indirect evidence, almost all of which the task force chose to ignore.

As it happens, a comprehensive survey of that evidence, and of the objections to it, appeared this past June in the journal *Psychology, Public Policy and Law*. There, J. Philippe Rushton and Arthur Jensen co-authored a 60-page article titled "Thirty Years of Research on Race Differences in Cognitive Ability." It incorporates studies of East Asians as well as blacks and whites and concludes that the source of the black-white-Asian difference is 50% to 80% genetic. The same issue of the journal includes four commentaries, three of them written by prominent scholars who oppose the idea that any part of the black-white difference is genetic. Thus, in one place, you can examine the strongest arguments that each side in the debate can bring to bear.

Messrs. Rushton and Jensen base their conclusion on 10 categories of evidence that are consistent with a model in which both environment and genes cause the black-white difference and inconsistent with a model that requires no genetic contribution. I will not try to review their argument here, or the critiques of it. All of the contributions can be found on the Internet, and can be understood by readers with a grasp of basic statistical concepts.

For those who consider it important to know what percentage of the IQ difference is genetic, a methodology that would do the job is now available. In the United States, few people classified as black are actually of 100% African descent (the average American black is thought to be about 20% white). To the extent that genes play a role, IQ will vary by racial admixture. In the past, studies that have attempted to test this hypothesis have had no accurate way to measure the degree of admixture, and the results have been accordingly muddled. The recent advances in using genetic markers solve that problem. Take a large sample of racially diverse people, give them a good IQ test, and then use genetic markers to create a variable that no longer classifies people as "white" or "black," but along a continuum. Analyze the variation in IQ scores according to that continuum. The results would be close to dispositive.

None of this is important for social policy, however, where the issue is not the source of the difference but its intractability. Much of the evidence reviewed by Messrs. Rushton and Jensen bears on what we can expect about future changes in the black-white IQ difference. My own thinking on this issue is shaped by the relationship of the difference to a factor I have already mentioned-- "*g*"--and to the developing evidence for *g*'s biological basis.



When you compare black and white mean scores on a battery of subtests, you do not find a uniform set of differences; nor do you find a random assortment. The size of the difference varies systematically by type of subtest. Asked to predict which subtests show the largest difference, most people will think first of ones that have the most cultural content and are the most sensitive to good schooling. But this natural expectation is wrong. Some of the largest differences are found on subtests that have little or no cultural content, such as ones based on abstract designs.

As long ago as 1927, Charles Spearman, the pioneer psychometrician who discovered g , proposed a hypothesis to explain the pattern: the size of the black-white difference would be "most marked in just those [subtests] which are known to be saturated with g ." In other words, Spearman conjectured that the black-white difference would be greatest on tests that were the purest measures of intelligence, as opposed to tests of knowledge or memory.

A concrete example illustrates how Spearman's hypothesis works. Two items in the Wechsler and Stanford-Binet IQ tests are known as "forward digit span" and "backward digit span." In the forward version, the subject repeats a random sequence of one-digit numbers given by the examiner, starting with two digits and adding another with each iteration. The subject's score is the number of digits that he can repeat without error on two consecutive trials. Digits-backward works exactly the same way except that the digits must be repeated in the opposite order.

Digits-backward is much more g -loaded than digits-forward. Try it yourself and you will see why. Digits-forward is a straightforward matter of short-term memory. Digits-backward makes your brain work much harder.

The black-white difference in digits-backward is about twice as large as the difference in digits-forward. It is a clean example of an effect that resists cultural explanation. It cannot be explained by differential educational attainment, income or any other socioeconomic factor. Parenting style is irrelevant. Reluctance to "act white" is irrelevant. Motivation is irrelevant. There is no way that any of these variables could systematically encourage black performance in digits-forward while depressing it in digits-backward in the same test at the same time with the same examiner in the same setting.

In 1980, Arthur Jensen began a research program for testing Spearman's hypothesis. In his book "The g Factor" (1998), he summarized the results from 17 independent sets of data, derived from 149 psychometric tests. They consistently supported Spearman's hypothesis. Subsequent work has added still more evidence. Debate continues about what the correlation between g -loadings and the size of the black-white difference means, but the core of Spearman's original conjecture, that a sizable correlation would be found to exist, has been confirmed.

During the same years that Mr. Jensen was investigating Spearman's hypothesis, progress was also being made in understanding g . For decades, psychometricians had tried to make g go away. Confident that intelligence must be more complicated than a single factor, they strove to replace g with measures of uncorrelated mental skills. They thereby made valuable contributions to our understanding of intelligence, which really does manifest itself in different ways and with different profiles, but getting rid of g proved impossible. No matter how the data were analyzed, a single factor kept dominating the results.

By the 1980s, the robustness and value of g as an explanatory construct were broadly

accepted among psychometricians, but little was known about its physiological basis. As of 2005, we know much more. It is now established that g is by far the most heritable component of IQ. A variety of studies have found correlations between g and physiological phenomena such as brain-evoked potentials, brain pH levels, brain glucose metabolism, nerve-conduction velocity and reaction time. Most recently, it has been determined that a highly significant relationship exists between g and the volume of gray matter in specific areas of the frontal cortex, and that the magnitude of the volume is under tight genetic control. In short, we now know that g captures something in the biology of the brain.

So Spearman's basic conjecture was correct--the size of the black-white difference and g -loadings are correlated--and g represents a biologically grounded and highly heritable cognitive resource. When those two observations are put together, a number of characteristics of the black-white difference become predictable, correspond with phenomena we have observed in data, and give us reason to think that not much will change in the years to come.



One implication is that black-white convergence on test scores will be greatest on tests that are least g -loaded. Literacy is the obvious example: People with a wide range of IQs can be taught to read competently, and it is the reading test of the NAEP in which convergence has reached its closest point (0.55 standard deviation in the 1988 test). More broadly, the confirmation of Spearman's hypothesis explains why the convergence that has occurred on academic achievement tests has not been matched on IQ tests.

A related implication is that the source of the black-white difference lies in skills that are hardest to change. Being able to repeat many digits backward has no value in itself. It points to a valuable underlying mental ability, in the same way that percentage of fast-twitch muscle fibers points to an underlying athletic ability. If you were to practice reciting digits backward for a few days, you could increase your score somewhat, just as training can improve your running speed somewhat. But in neither case will you have improved the underlying ability. As far as anyone knows, g itself cannot be coached.

The third implication is that the "Flynn effect" will not close the black-white difference. I am referring here to the secular increase in IQ scores over time, brought to public attention by James Flynn. The Flynn effect has been taken as a reason for thinking that the black-white difference is temporary: If IQ scores are so malleable that they can rise steadily for several decades, why should not the black-white difference be malleable as well?

But as the Flynn effect has been studied over the past decade, the evidence has grown, and now seems persuasive, that the increases in IQ scores do not represent significant increases in g . What the increases do represent--whether increases in specific mental skills or merely increased test sophistication--is still being debated. But if the black-white difference is concentrated in g and if the Flynn effect does not consist of increases in g ,

the Flynn effect will not do much to close the gap. A 2004 study by Dutch scholars tested this question directly. Examining five large databases, the authors concluded that "the nature of the Flynn effect is qualitatively different from the nature of black-white differences in the United States," and that "the implications of the Flynn effect for black-white differences appear small."

These observations represent my reading of a body of evidence that is incomplete, and they will surely have to be modified as we learn more. But taking the story of the black-white IQ difference as a whole, I submit that we know two facts beyond much doubt. First, the conventional environmental explanation of the black-white difference is inadequate. Poverty, bad schools and racism, which seem such obvious culprits, do not explain it. Insofar as the environment *is* the cause, it is not the sort of environment we know how to change, and we have tried every practical remedy that anyone has been able to think of. Second, regardless of one's reading of the competing arguments, we are left with an IQ difference that has, at best, narrowed by only a few points over the last century. I can find nothing in the history of this difference, or in what we have learned about its causes over the last ten years, to suggest that any faster change is in our future.



Elites throughout the West are living a lie, basing the futures of their societies on the assumption that all groups of people are equal in all respects. Lie is a strong word, but justified. It is a lie because so many elite politicians who profess to believe it in public do not believe it in private. It is a lie because so many elite scholars choose to ignore what is already known and choose not to inquire into what they suspect. We enable ourselves to continue to live the lie by establishing a taboo against discussion of group differences.

The taboo is not perfect--otherwise, I would not have been able to document this essay--but it is powerful. Witness how few of Harvard's faculty who understood the state of knowledge about sex differences were willing to speak out during the Summers affair. In the public-policy debate, witness the contorted ways in which even the opponents of policies like affirmative action frame their arguments so that no one can accuse them of saying that women are different from men or blacks from whites. Witness the unwillingness of the mainstream media to discuss group differences without assuring readers that the differences will disappear when the world becomes a better place.

The taboo arises from an admirable idealism about human equality. If it did no harm, or if the harm it did were minor, there would be no need to write about it. But taboos have consequences.

The nature of many of the consequences must be a matter of conjecture because people are so fearful of exploring them. Consider an observation furtively voiced by many who interact with civil servants: that government is riddled with people who have been promoted to their level of incompetence because of pressure to have a staff with the

correct sex and ethnicity in the correct proportions and positions. Are these just anecdotes? Or should we be worrying about the effects of affirmative action on the quality of government services? It would be helpful to know the answers, but we will not so long as the taboo against talking about group difference prevails.

How much damage has the taboo done to the education of children? Christina Hoff Sommers has argued that willed blindness to the different developmental patterns of boys and girls has led many educators to see boys as aberrational and girls as the norm, with pervasive damage to the way our elementary and secondary schools are run. Is she right? Few have been willing to pursue the issue lest they be required to talk about innate group differences. Similar questions can be asked about the damage done to medical care, whose practitioners have only recently begun to acknowledge the ways in which ethnic groups respond differently to certain drugs.

How much damage has the taboo done to our understanding of America's social problems? The part played by sexism in creating the ratio of males to females on mathematics faculties is not the ratio we observe but what remains after adjustment for male-female differences in high-end mathematical ability. The part played by racism in creating different outcomes in black and white poverty, crime and illegitimacy is not the raw disparity we observe but what remains after controlling for group characteristics. For some outcomes, sex or race differences nearly disappear after a proper analysis is done. For others, a large residual difference remains. In either case, open discussion of group differences would give us a better grasp on where to look for causes and solutions.



What good can come of raising this divisive topic? The honest answer is that no one knows for sure. What we do know is that the taboo has crippled our ability to explore almost any topic that involves the different ways in which groups of people respond to the world around them--which means almost every political, social or economic topic of any complexity.

Thus my modest recommendation, requiring no change in laws or regulations, just a little more gumption. Let us start talking about group differences openly--all sorts of group differences, from the visuospatial skills of men and women to the vivaciousness of Italians and Scots. Let us talk about the nature of the manly versus the womanly virtues. About differences between Russians and Chinese that might affect their adoption of capitalism. About differences between Arabs and Europeans that might affect the assimilation of Arab immigrants into European democracies. About differences between the poor and nonpoor that could inform policy for reducing poverty.

Even to begin listing the topics that could be enriched by an inquiry into the nature of group differences is to reveal how stifled today's conversation is. Besides liberating that conversation, an open and undefensive discussion would puncture the irrational fear of the male-female and black-white differences I have surveyed here. We would be free to talk about other sexual and racial differences as well, many of which favor women and

blacks, and none of which is large enough to frighten anyone who looks at them dispassionately.

Talking about group differences does not require any of us to change our politics. For every implication that the right might seize upon (affirmative-action quotas are ill-conceived), another gives fodder to the left (innate group differences help rationalize compensatory redistribution by the state). But if we do not need to change our politics, talking about group differences obligates all of us to renew our commitment to the ideal of equality that Thomas Jefferson had in mind when he wrote as a self-evident truth that all men are created equal. Steven Pinker put that ideal in today's language in "The Blank Slate," writing that "equality is not the empirical claim that all groups of humans are interchangeable; it is the moral principle that individuals should not be judged or constrained by the average properties of their group."

Nothing in this essay implies that this moral principle has already been realized or that we are powerless to make progress. In elementary and secondary education, many outcomes are tractable even if group differences in ability remain unchanged. Dropout rates, literacy and numeracy are all tractable. School discipline, teacher performance and the quality of the curriculum are tractable. Academic performance within a given IQ range is tractable. The existence of group differences need not and should not discourage attempts to improve schooling for millions of American children who are now getting bad educations.

In university education and in the world of work, overall openness of opportunity has been transformed for the better over the past half-century. But the policies we now have in place are impeding, not facilitating, further progress. Creating double standards for physically demanding jobs so that women can qualify ensures that men in those jobs will never see women as their equals. In universities, affirmative action ensures that the black-white difference in IQ in the population at large is brought onto the campus and made visible to every student. The intentions of their designers notwithstanding, today's policies are perfectly fashioned to create separation, condescension and resentment--and so they have done.

The world need not be that way. Any university or employer that genuinely applied a single set of standards for hiring, firing, admitting and promoting would find that performance really is distributed indistinguishably across different groups. But getting to that point nationwide will require us to jettison an apparatus of laws, regulations and bureaucracies that has been 40 years in the making. That will not happen until the conversation has opened up. So let us take one step at a time. Let us stop being afraid of data that tell us a story we do not want to hear, stop the name-calling, stop the denial and start facing reality.

Mr. Murray is W.H. Brady Scholar in Freedom and Culture at the American Enterprise Institute. This article appears in the September issue of Commentary. A fully annotated version, which includes extensive supplementary material, is available [here](#).

Copyright © 2005 Dow Jones & Company, Inc. All Rights Reserved.

[PRINT WINDOW](#) [CLOSE WINDOW](#)