

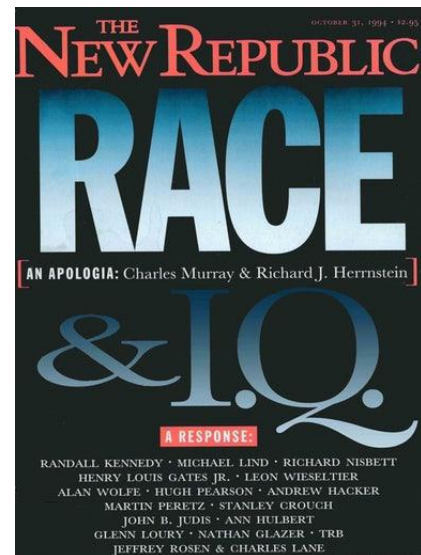
[Race, Genes and I.Q. — An Apologia*](#)

Charles Murray and Richard J. Herrnstein

The New Republic, Vol. 211 Issue 18, October 31, 1994

I.

The private dialogue about race in America is far different from the public one, and we are not referring just to discussions among white rednecks. Our impression is that the private attitudes of white elites toward blacks is strained far beyond any public acknowledgment, that hostility is not uncommon and that a key part of the strain is a growing suspicion that fundamental racial differences are implicated in the social and economic gap that continues to separate blacks and whites, especially alleged genetic differences in intelligence.



We say “our impression” because we have been in a unique position to gather impressions. Since the beginning of 1990, we have been writing a book about differences in intellectual capacity among people and groups and what those differences mean for America’s future. As authors do, we have gotten into numberless conversations that begin, “What are you working on now?” Our interlocutors have included scholars at the top-ranked universities and think tanks, journalists, high public officials, lawyers, financiers and corporate executives. In the aggregate, they have split about evenly between left and right of the political center.

With rare exceptions, these people have shared one thing besides their success. As soon as the subject turned to the question of I.Q., they focused on whether there was any genetic race differences in intelligence. And they tended to be scared stiff about the answer. This experience has led us to be scared as well, about the consequences of ignorance. We have been asked whether the question of racial genetic differences in intelligence should even be raised in polite society. We believe there’s no alternative. A taboo issue, filled with potential for hurt and anger, lurks just beneath the surface of American life. It is essential that people begin to talk about this in the open. Because raising this question at all provokes a host of fears, it is worth stating at the outset a clear conclusion of our research: the fascination with race, I.Q. and genes is misbegotten. There are all sorts of things to be worried about regarding intelligence and American life, and even regarding intelligence and ethnicity. But genetics isn’t one of them.

II.

First, the evidence, beginning with this furiously denied fact: intelligence is a useful construct. Among the experts, it is by now beyond much technical dispute that there is such a thing as a general factor of cognitive ability on which human beings differ and that this general factor is measured reasonably well by a variety of standardized tests, best of all by I.Q. tests designed for that purpose. These points are no longer the topic of much new work in the technical journals because most of the questions about them have been answered.

* This article was originally published in the October 31, 1994, issue of *The New Republic*. Since many staffers at the time objected to its publication, this excerpt of *The Bell Curve* was published alongside [a raft of articles](#) condemning it.

Intelligence as measured by I.Q. tests is predictive of many educational, economic and social outcomes. In America today, you are much better off knowing a child's I.Q. score than her parents' income or education if you want to predict whether she will drop out of high school, for example. If you are an employer trying to predict an applicant's job productivity and are given a choice of just one item of information, you are usually better off asking for an I.Q. score than a resume, college transcript, letter of recommendation or even a job interview. These statements hold true for whites, blacks, Asians and Latinos alike.

This is not to say that I.Q. is destiny—in each of these instances, I.Q. is merely a better predictor than the alternatives, not even close to a perfect one. But it should be stated that the pariah status of intelligence as a construct and I.Q. as its measure for the past three decades has been a function of political fashion, not science.

Ethnic differences in measured cognitive ability have been found since intelligence tests were invented. The battle over the meaning of these differences is largely responsible for today's controversy over intelligence testing itself. The first thing to remember is that the differences among individuals are far greater than the differences among groups. If all the ethnic differences in intelligence evaporated overnight, most of the intellectual variation in America would endure. The remaining inequality would still strain the political process, because differences in cognitive ability are problematic even in ethnically homogeneous societies.

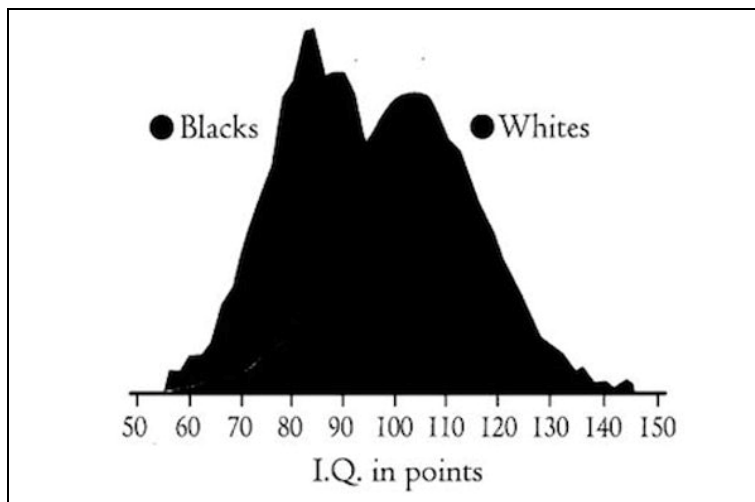
Even using the word "race" is problematic, which is why we use the word ethnicity as well as race in this article. What does it mean to be "black" in America, in racial terms, when the word black (or African American) can be used for people whose ancestry is more European than African? How are we to classify a person whose parents hail from Panama but whose ancestry is predominantly African? Is he Latino? Black? The rule we follow here is a simple one: to classify people according to the way they classify themselves.

III.

We might start with a common question in America these days: Do Asians have higher I.Q.s than whites? The answer is probably yes, if Asian refers to the Japanese and Chinese (and perhaps also Koreans), whom we will refer to here as East Asians. How much higher is still unclear. The best tests of this have involved identical I.Q. tests given to populations that are comparable except for race. In one test, samples of American, British and Japanese students aged 13 to 15 were given a test of abstract reasoning and spatial relations. The U.S. and U.K. samples had scores within a point of the standardized mean of 100 on both the abstract and spatial relations parts of the test; the Japanese scored 104.5 on the test for abstract reasoning and 114 on the test for spatial relations—a large difference, amounting to a gap similar to the one found by another leading researcher for Asians in America. In a second set of studies, 9-year-olds in Japan, Hong Kong and Britain, drawn from comparable socioeconomic populations, were administered the Ravens Standard Progressive Matrices. The children from Hong Kong averaged 113; from Japan, 110; and from Britain, 100.

Not everyone accepts that the East Asian-white difference exists. Another set of studies gave a battery of mental tests to elementary school children in Japan, Taiwan and Minneapolis, Minnesota. The key difference between this study and the other two was that the children were matched carefully on many socioeconomic and demographic variables. No significant difference in overall I.Q. was found, and the authors concluded that "this study offers no support for the argument that there are differences in the general cognitive functioning of Chinese, Japanese and American children."

Where does this leave us? The parties in the debate are often confident, and present in their articles are many flat statements that an overall East Asian-white I.Q. difference does, or does not, exist. In our judgment, the balance of the evidence supports the notion that the overall East Asian mean is higher than the white mean. Three I.Q. points most resembles a consensus, tentative though it still is. East Asians have a greater advantage in a particular kind of nonverbal intelligence.



Black & White I.Q. Distribution for populations of equal size

The issues become far more fraught, however, in determining the answer to the question: Do African Americans score differently from whites on standardized tests of cognitive ability? If the samples are chosen to be representative of the American population, the answer has been yes for every known test of cognitive ability that meets basic psychometric standards. The answer is also yes for almost all studies in which the black and white samples are matched on some special characteristics—juvenile delinquents, for example, or graduate students—but there are exceptions.

How large is the black-white difference? The usual answer is what statisticians call one standard deviation. In discussing I.Q. tests, for example, the black mean is commonly given as 85, the white mean as 100 and the standard deviation as fifteen points. But the differences observed in any given study seldom conform exactly to one standard deviation. In 156 American studies conducted during this century that have reported the I.Q. means of a black sample and a white sample, and that meet basic requirements of interpretability, the mean black-white difference is 1.1 standard deviations, or about sixteen I.Q. points.

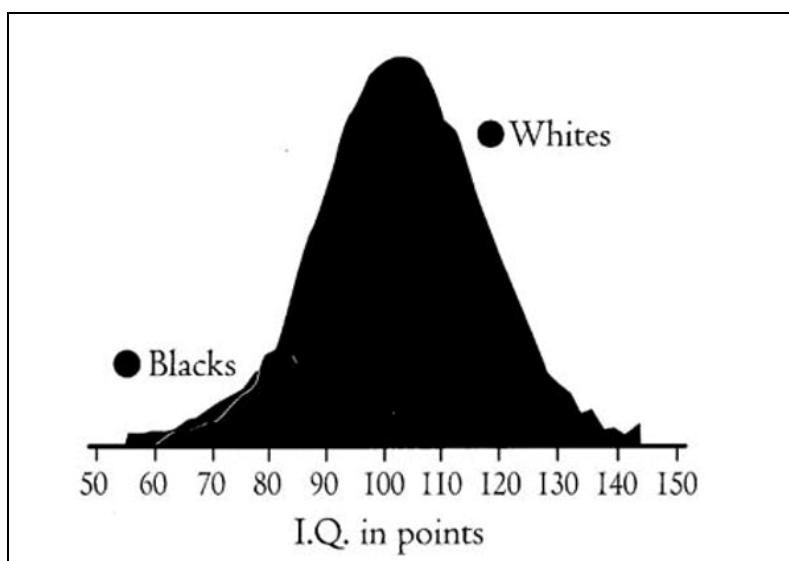
More rigorous selection criteria do not diminish the size of the gap. For example, with tests given outside the South only after 1960, when people were increasingly sensitized to racial issues, the number of studies is reduced to twenty-four, but the mean difference is still 1.1 standard deviations. The National Longitudinal Survey of Youth (NLSY) administered an I.Q. test in 1980 to by far the largest and most carefully selected national sample (6,502 whites, 3,022 blacks) and found a difference of 1.2 standard deviations.

Evidence from the SAT, the ACT and the National Assessment of Educational Progress gives reason to think that the black-white I.Q. difference has shrunk by perhaps three I.Q. points in the last twenty years. Almost all the improvement came in the low end, however, progress has stalled for several years and the most direct evidence, from I.Q. tests of the next generation in the nlsy, points to a widening black-white gap rather than a shrinking one.

It is important to understand that even a difference of 1.2 standard deviations means considerable overlap in the cognitive ability distribution for blacks and whites, as shown for the NLSY population in the figure on page 28. For any equal number of blacks and whites, a large proportion have I.Q.s that can be matched up. For that matter, millions of blacks have higher I.Q.s than the average white. Tens of thousands have I.Q.s that put them in the top few percentiles of the white distribution. It should be no surprise to see (as everyone does every day) African Americans functioning at high levels in every intellectually challenging field. This is the distribution to keep in mind whenever thinking about individuals.

But an additional complication must be taken into account: in the United States, there are about six whites for every black. This means that the I.Q. overlap of the two populations as they actually exist in the United States looks very different from the overlap in the figure on page 28. The figure above presents the same data from the NLSY when the distributions are shown in proportion to the actual population of young people in the NLSY. This figure shows why a black-white difference can be problematic to society as a whole. At the lower end of the I.Q. range, there are about equal numbers of blacks and whites. But throughout the upper half of the range, the disproportions between the number of whites and blacks at any given I.Q. level are huge. To the extent that the difference represents an authentic difference in cognitive functioning, the social consequences are huge as well. But is the difference authentic? Is it, for example, attributable to cultural bias or other artifacts of the test? There are several ways of assessing this. We'll go through them one by one.

External evidence of bias. Tests are used to predict things—most commonly, to predict performance in school or on the job. The ability of a test to predict is known as its validity. A test with high validity predicts accurately; a test with poor validity makes many mistakes. Now suppose that a test's validity differs for the members of two groups. To use a concrete example: the SAT is used as a tool in college admissions because it has a certain validity in predicting college performance. If the SAT is biased against blacks, it will underpredict their college performance. If tests were biased in this way, blacks as a group would do better in college than the admissions office expected based just on their SATs. It would be as if the test underestimated the "true" SAT score of the blacks, so the natural remedy for this would be to compensate the black applicants by, for example, adding the appropriate number of points to their scores.



Black & White I.Q. Distribution Proportional to the ethnic composition of the U.S.

Predictive bias can work in another way, as when the test is simply less reliable—that is, less accurate—for blacks than for whites. Suppose a test used to select police sergeants is more accurate in predicting the performance of white candidates who become sergeants than in predicting the performance of black sergeants. It doesn't underpredict for blacks, but rather fails to predict at all (or predicts less accurately). In these cases, the natural remedy would be to give less weight to the test scores of blacks than to those of whites.

The key concept for both types of bias is the same: a test biased against blacks does not predict black performance in the real world in the same way that it predicts white performance in the real world. The evidence of bias is external in the sense that it shows up in differing validities for blacks and whites. External evidence of bias has been sought in hundreds of studies. It has been evaluated relative to performance in elementary school, in the university, in the military, in unskilled and skilled jobs, in the professions. Overwhelmingly, the evidence is that the standardized tests used to help make school and job decisions do not underpredict black performance. Nor does the expert community find any other systematic difference in the predictive accuracy of tests for blacks and whites.

Internal evidence of bias. The most common charges of cultural bias involve the putative cultural loading of items in a test. Here is an SAT analogy item that has become famous as an example of cultural bias:

RUNNER: MARATHON (A) envoy: embassy (B) martyr: massacre (C) oarsman: regatta (D) referee: tournament (E) horse: stable

The is "oarsman: regatta"—fairly easy if you know what both a marathon and a regatta are, a matter of guesswork otherwise. How would a black youngster from the inner city ever have heard of a regatta? Many view such items as proof that the tests must be biased against people from disadvantaged backgrounds. "Clearly," writes a critic of testing, citing this example, "this item does not measure students' aptitude or logical reasoning ability, but knowledge of upper-middle-class recreational activity." In the language of psychometrics, this is called internal evidence of bias.

The hypothesis of bias again lends itself to direct examination. In effect, the SAT critic is saying that culturally loaded items are producing at least some of the black-white difference. Get rid of such items, and the gap will narrow. Is he correct? When we look at the results for items that have answers such as "oarsman: regatta" and the results for items that seem to be empty of any cultural information (repeating a sequence of numbers, for example), are there any differences?

The technical literature is again clear. In study after study of the leading tests, the idea that the black-white difference is caused by questions with cultural content has been contradicted by the facts. Items that the average white test-taker finds easy relative to other items, the average black test-taker does, too; the same is true for items that the average white and black find difficult. Inasmuch as whites and blacks have different overall scores on the average, it follows that a smaller proportion of blacks get right answers for either easy or hard items, but the order of difficulty is virtually the same in each racial group. How can this be? The explanation is complicated and goes deep into the reasons why a test item is "good" or "bad" in measuring intelligence. Here, we restrict ourselves to the conclusion: The black-white difference is generally wider on items that appear to be culturally neutral than on items that appear to be culturally loaded. We italicize this point because it is so well established empirically yet comes as such a surprise to most people who are new to this topic.

Motivation to try. Suppose the nature of cultural bias does not lie in predictive validity or in the content of the items but in what might be called “test willingness.” A typical black youngster, it is hypothesized, comes to such tests with a mindset different from the white subject’s. He is less attuned to testing situations (from one point of view), or less inclined to put up with such nonsense (from another). Perhaps he just doesn’t give a damn, since he has no hopes of going to college or otherwise benefiting from a good test score. Perhaps he figures that the test is biased against him anyway, so what’s the point. Perhaps he consciously refuses to put forth his best effort because of the peer pressure against “acting white” in some inner-city schools.

The studies that have attempted to measure motivation in such situations generally have found that blacks are at least as motivated as whites. But these are not wholly convincing, for why shouldn’t the measures of motivation be just as inaccurate as the measures of cognitive ability are alleged to be? Analysis of internal characteristics of the tests once again offers the best leverage in examining this broad hypothesis. Here, we will offer just one example involving the “digit span” subtest, part of the widely used Wechsler intelligence tests. It has two forms: forward digit span, in which the subject tries to repeat a sequence of numbers in the order read to him, and backward digit span, in which the subject tries to repeat the sequence of numbers backward. The test is simple, uses numbers familiar to everyone and calls on no cultural information besides numbers. The digit span is informative regarding test motivation not just because of the low cultural loading of the items but because the backward form is a far better measure of “g,” the psychometrician’s shorthand for the general intelligence factor that I.Q. tests try to measure. The reason that the backward form is a better measure of g is that reversing the numbers is mentally more demanding than repeating them in the heard order, as you can determine for yourself by a little self-testing.

The two parts of the subtest have identical content. They occur at the same time during the test. Each subject does both. But in most studies the black-white difference is about twice as great on backward digits as on forward digits. The question then arises: How can lack of motivation (or test willingness) explain the difference in performance on the two parts of the same subtest?

This still leaves another obvious question: Are the differences in overall black and white test scores attributable to differences in socioeconomic status? This question has two different answers depending on how the question is understood, and confusion is rampant. There are two essential answers and two associated rationales.

First version: If you extract the effects of socioeconomic class, what happens to the magnitude of the black-white difference? Blacks are disproportionately in the lower socioeconomic classes, and class is known to be associated with I.Q. Therefore, many people suggest, part of what appears to be an ethnic difference in I.Q. scores is actually a socioeconomic difference. The answer to this version of the question is that the size of the gap shrinks when socioeconomic status is statistically extracted. The NLSY gives a result typical of such analyses. The black-white difference in the NLSY is 1.2. In a regression equation in which both race and socioeconomic background are entered, the difference between whites and blacks shrinks to less than .8 standard deviation. Socioeconomic status explains 37 percent of the original black-white difference. This relationship is in line with the results from many other studies.

The difficulty comes in interpreting what it means to “control” for socioeconomic status. Matching the status of the groups is usually justified on the grounds that the scores people earn are caused to some extent by their socioeconomic status, so if we want to see the “real”

or “authentic” difference between them, the contribution of status must be excluded. The trouble is that socioeconomic status is also a result of intelligence, as people of high and low cognitive ability move to high and low places in the class structure. The reason parents have high or low socioeconomic status is in part a function of their intelligence, and their intelligence also affects the I.Q. of the children via both genes and environment.

Because of these relationships, “controlling” for socioeconomic status in racial comparisons is guaranteed to reduce I.Q. differences in the same way that choosing black and white samples from a school for the intellectually gifted is guaranteed to reduce I.Q. differences (assuming race-blind admissions standards). These complications aside, a reasonable rule of thumb is that controlling for socioeconomic status reduces the overall black-white difference by about one-third.

Second version: As blacks move up the socioeconomic ladder, do the differences with whites of similar socioeconomic status diminish? The first version of the SES/I.Q. question referred to the overall score of a population of blacks and whites. The second version concentrates on the black-white difference within socioeconomic classes. The rationale goes like this: blacks score lower on average because they are socioeconomically at a disadvantage. This disadvantage should most seriously handicap children in the lower socioeconomic classes, who suffer from greater barriers to education and job advancement than do children in the middle and upper classes. As blacks advance up the socioeconomic ladder, their children, less exposed to these barriers, will do better and, by extension, close the gap with white children of their class.

This expectation is not borne out by the data. A good way to illustrate this is to use an index of parental ses based on their education, income and occupation and to match it against the mean I.Q. score, as shown in the figure on page 32. I.Q. scores increase with economic status for both races. But as the figure shows, the magnitude of the black-white difference in standard deviations does not decrease. Indeed, it gets larger as people move up from the very bottom of the socioeconomic ladder. The pattern shown in the figure is consistent with many other major studies, except that the gap flattens out. In other studies, the gap has continued to increase throughout the range of socioeconomic status.

IV.

This brings us to the flashpoint of intelligence as a public topic: the question of genetic differences between the races. Expert opinion, when it is expressed at all, diverges widely. In the 1980s Mark Snyderman, a psychologist, and Stanley Rothman, a political scientist, sent a questionnaire to a broad sample of 1,020 scholars, mostly academicians, whose specialties give them reason to be knowledgeable about I.Q. Among other questions, they asked, “Which of the following best characterizes your opinion of the heritability of the black-white difference in I.Q.?” The answers were divided as follows: The difference is entirely due to environmental variation: 15 percent. The difference is entirely due to genetic variation: 1 percent. The difference is a product of both genetic and environmental variation: 45 percent. The data are insufficient to support any reasonable opinion: 24 percent. No response: 14 percent.

This pretty well sums up the professional judgment on the matter. But it doesn’t explain anything about the environment/genetic debate as it has played out in the profession and in the general public. And the question, of course, is fascinating. So what could help us understand the connection between heritability and group differences? A good place to start is by correcting a common confusion about the role of genes in individuals and in groups.

Most scholars accept that I.Q. in the human species as a whole is substantially heritable, somewhere between 40 percent and 80 percent, meaning that much of the observed variation in I.Q. is genetic. And yet this information tells us nothing for sure about the origin of the differences between groups of humans in measured intelligence. This point is so basic, and so misunderstood, that it deserves emphasis: that a trait is genetically transmitted in a population does not mean that group differences in that trait are also genetic in origin. Anyone who doubts this assertion may take two handfuls of genetically identical seed corn and plant one handful in Iowa, the other in the Mojave Desert, and let nature (i.e., the environment) take its course. The seeds will grow in Iowa, not in the Mojave, and the result will have nothing to do with genetic differences.

The environment for American blacks has been closer to the Mojave and the environment for American whites has been closer to Iowa. We may apply this general observation to the available data and see where the results lead. Suppose that all the observed ethnic differences in tested intelligence originate in some mysterious environmental differences—mysterious, because we know from material already presented that socioeconomic factors cannot be much of the explanation. We further stipulate that one standard deviation (fifteen I.Q. points) separates American blacks and whites and that one-fifth of a standard deviation (three I.Q. points) separates East Asians and whites. Finally, we assume that I.Q. is 60 percent heritable (a middle-ground estimate). Given these parameters, how different would the environments for the three groups have to be in order to explain the observed difference in these scores?

The observed ethnic differences in I.Q. could be explained solely by the environment if the mean environment of whites is 1.58 standard deviations better than the mean environment of blacks and .32 standard deviation worse than the mean environment for East Asians, when environments are measured along the continuum of their capacity to nurture intelligence. Let's state these conclusions in percentile terms: the average environment of blacks would have to be at the sixth percentile of the distribution of environments among whites and the average environment of East Asians would have to be at the sixty-third percentile of environments among whites for the racial differences to be entirely environmental.

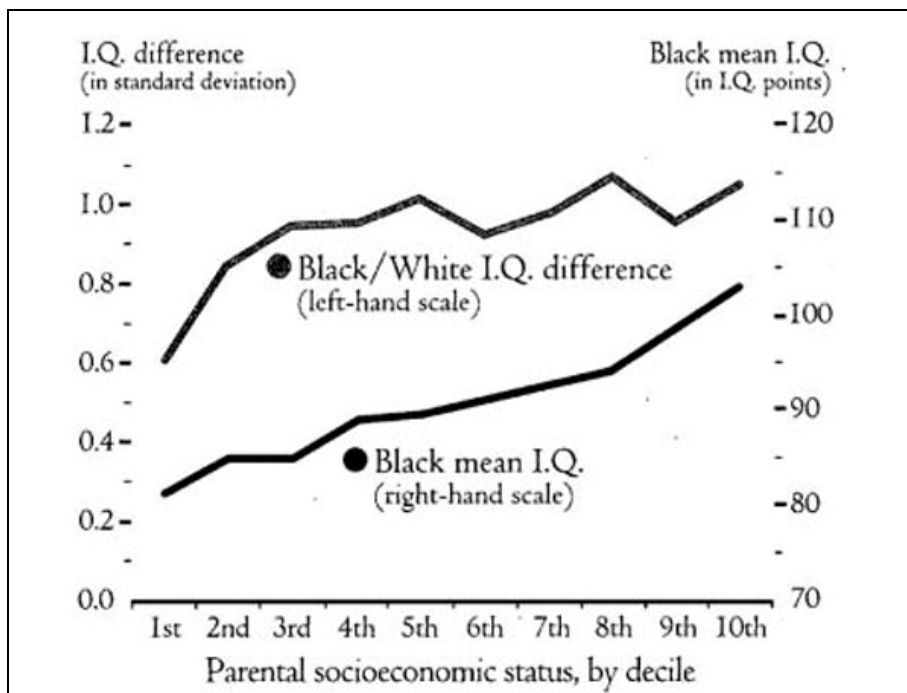
Environmental differences of this magnitude and pattern are wildly out of line with all objective measures of the differences in black, Asian and white environments. Recall further that the black-white difference is smallest at the lowest socioeconomic levels. Why, if the black-white difference is entirely environmental, should the advantage of the "white" environment compared to the "black" be greater among the better-off and better-educated blacks and whites? We have not been able to think of a plausible reason. Can you? An appeal to the effects of racism to explain ethnic differences also requires explaining why environments poisoned by discrimination and racism for some other groups—against the Chinese or the Jews in some regions of America for example—have left them with higher scores than the national average.

However discomfiting it may be to consider it, there are reasons to suspect genetic considerations are involved. The evidence is circumstantial, but provocative. For example, ethnicities differ not just in average scores but in the profile of intellectual capacities. A full-scale I.Q. score is the aggregate of many subtests. There are thirteen of them in the Wechsler Intelligence Scale for Children, for example. The most basic division of the subtests is into a verbal I.Q. and a performance I.Q. In white samples the verbal and performance I.Q. subscores tend to have about the same mean, because I.Q. tests have been standardized on predominantly white populations. But individuals can have imbalances between these two I.Q.s. People with high verbal abilities are likely to do well with words and logic. In school

they excel in history and literature; in choosing a career to draw on those talents, they tend to choose law or journalism or advertising or politics. In contrast, people with high performance I.Q.s—or, using a more descriptive phrase, “visuospatial abilities”—are likely to do well in the physical and biological sciences, mathematics, engineering or other subjects that demand mental manipulation in the three physical dimensions or the more numerous dimensions of mathematics.

East Asians living overseas score about the same or slightly lower than whites on verbal I.Q. and substantially higher on visuospatial I.Q. Even in the rare studies that have found overall Japanese or Chinese I.Q.s no higher than white I.Q.s, the discrepancy between verbal and visuospatial I.Q. persists. For Japanese living in Asia, a 1987 review of the literature demonstrated without much question that the verbal-visuospatial difference persists even in examinations that have been thoroughly adapted to the Japanese language and, indeed, in tests developed by the Japanese themselves. A study of a small sample of Korean infants adopted into white families in Belgium found the familiar elevated visuospatial scores.

This finding has an echo in the United States, where Asian American students abound in science subjects, in engineering and in medical schools, but are scarce in law schools and graduate programs in the humanities and social sciences. Is this just a matter of parental pressures or of Asian immigrants uncomfortable with English? The same pattern of subtest scores is found in Inuits and American Indians (both of Asian origin) and in fully assimilated second- and third-generation Asian Americans. Any simple socioeconomic, cultural or linguistic explanation is out of the question, given the diversity of living conditions, native languages, educational systems and cultural practices experienced by these groups and by East Asians living in Asia. Their common genetic history cannot plausibly be dismissed as irrelevant.



Black I.Q. scores rise with socioeconomic status, but the Black/White difference remains.

Turning now to blacks and whites (using these terms to refer exclusively to Americans), ability profiles also have been important in understanding the nature, and possible genetic component, of group differences. The argument has been developing around what is known as Spearman's hypothesis. This hypothesis says that if the black-white difference on test scores reflects a real underlying difference in general mental ability (g), then the size of the

black-white difference will be related to the degree to which the test is saturated with g. In other words, the better a test measures g, the larger the black-white difference will be.

By now, Spearman's hypothesis has been borne out in fourteen major studies, and no appropriate data set has yet been found that contradicts Spearman's hypothesis. It should be noted that not all group differences behave similarly. For example, deaf children often get lower test scores than hearing children, but the size of the difference is not correlated positively with the test's loading on g. The phenomenon seems peculiarly concentrated in comparisons of ethnic groups. How does this bear on the genetic explanation of ethnic differences? In plain though somewhat imprecise language: the broadest conception of intelligence is embodied in g. At the same time, g typically has the highest heritability (higher than the other factors measured by I.Q. tests). As mental measurement focuses most specifically and reliably on g, the observed black-white mean difference in cognitive ability gets larger. This does not in itself demand a genetic explanation of the ethnic difference but, by asserting that "the better the test, the greater the ethnic difference," Spearman's hypothesis undercuts many of the environmental explanations of the difference that rely on the proposition (again, simplifying) that the apparent black-white difference is the result of bad tests, not good ones.

There are, of course, many arguments against such a genetic explanation. Many studies have shown that the disadvantaged environment of some blacks has depressed their test scores. In one study, in black families in rural Georgia, the elder sibling typically had a lower I.Q. than the younger. The larger the age difference is between the siblings, the larger is the difference in I.Q. The implication is that something in the rural Georgia environment was depressing the scores of black children as they grew older. In neither the white families of Georgia, nor white or black families in Berkeley, California, were there comparable signs of a depressive effect of the environment.

Another approach is to say that tests are artifacts of a culture, and a culture may not diffuse equally into every household and community. In a heterogeneous society, subcultures vary in ways that inevitably affect scores on I.Q. tests. Fewer books in the home mean less exposure to the material that a vocabulary subtest measures; the varying ways of socializing children may influence whether a child acquires the skills, or a desire for the skills, that tests test; the "common knowledge" that tests supposedly draw on may not be common in certain households and neighborhoods.

So far, this sounds like a standard argument about cultural bias, and yet it accepts the generalizations that we discussed earlier about internal evidence of bias. The supporters of this argument are not claiming that less exposure to books means that blacks score lower on vocabulary questions but do as well as whites on culture-free items. Rather, the effects of culture are more diffuse.

Furthermore, strong correlations between home or community life and I.Q. scores are readily found. In a study of 180 Latino and 180 non-Latino white elementary school children in Riverside, California, the researcher examined eight sociocultural variables: (1) mother's participation in formal organizations, (2) living in a segregated neighborhood, (3) home language level, (4) socioeconomic status based on occupation and education of head of household, (5) urbanization, (6) mother's achievement values, (7) home ownership, and (8) intact biological family. She then showed that once these sociocultural variables were taken into account, the remaining group and I.Q. differences among the children fell to near zero.

The problem with this procedure lies in determining what, in fact, these eight variables control for: cultural diffusion, or genetic sources of variation in intelligence as ordinarily understood? By so drastically extending the usual match for socioeconomic status, the possibility is that such studies demonstrate only that parents matched on I.Q. will produce children with similar I.Q.s—not a startling finding. Also, the data used for such studies continue to show the distinctive racial patterns in the subtests. Why should cultural diffusion manifest itself by differences in backward and forward digit span or in completely nonverbal items? If the role of European white cultural diffusion is so important in affecting black I.Q. scores, why is it so unimportant in affecting Asian I.Q. scores?

There are other arguments related to cultural bias. In the American context, Wade Boykin is one of the most prominent academic advocates of a distinctive black culture, arguing that nine interrelated dimensions put blacks at odds with the prevailing Eurocentric model. Among them are spirituality (blacks approach life as “essentially vitalistic rather than mechanistic, with the conviction that nonmaterial forces influence people’s everyday lives”); a belief in the harmony between humankind and nature; an emphasis on the importance of movement, rhythm, music and dance, “which are taken as central to psychological health”; personal styles that he characterizes as “verve” (high levels of stimulation and energy) and “affect” (emphasis on emotions and expressiveness); and “social time perspective,” which he defines as “an orientation in which time is treated as passing through a social space rather than a material one.” Such analyses purport to explain how large black-white differences in test scores could coexist with equal predictive validity of the test for such things as academic and job performance and yet still not be based on differences in “intelligence,” broadly defined, let alone genetic differences.

John Ogbu, a Berkeley anthropologist, has proposed a more specific version of this argument. He suggests that we look at the history of various minority groups to understand the sources of differing levels of intellectual attainment in America. He distinguishes three types of minorities: “autonomous minorities” such as the Amish, Jews and Mormons, who, while they may be victims of discrimination, are still within the cultural mainstream; “immigrant minorities,” such as the Chinese, Filipinos, Japanese and Koreans within the United States, who moved voluntarily to their new societies and, while they may begin in menial jobs, compare themselves favorably with their peers back in the home country; and, finally, “castelike minorities,” such as black Americans, who were involuntary immigrants or otherwise are consigned from birth to a distinctively lower place on the social ladder. Ogbu argues that the differences in test scores are an outcome of this historical distinction, pointing to a number of castes around the world—the untouchables in India, the Buraku in Japan and Oriental Jews in Israel—that have exhibited comparable problems in educational achievement despite being of the same racial group as the majority.

Indirect support for the proposition that the observed black-white difference could be the result of environmental factors is provided by the worldwide phenomenon of rising test scores. We call it “the Flynn effect” because of psychologist James Flynn’s pivotal role in focusing attention on it, but the phenomenon itself was identified in the 1930s when testers began to notice that I.Q. scores often rose with every successive year after a test was first standardized. For example, when the Stanford-Binet I.Q. was restandardized in the mid-1930s, it was observed that individuals earned lower I.Q.s on the new tests than they got on the Stanford-Binet that had been standardized in the mid-1910s; in other words, getting a score of 100 (the population average) was harder to do on the later test. This meant that the average person could answer more items on the old test than on the new test. Most of the change has been concentrated in the nonverbal portions of the tests.

The tendency for I.Q. scores to drift upward as a function of years since standardization has now been substantiated in many countries and on many I.Q. tests besides the Stanford-Binet. In some countries, the upward drift since World War II has been as much as a point per year for some spans of years. The national averages have in fact changed by amounts that are comparable to the fifteen or so I.Q. points separating whites and blacks in America. To put it another way, on the average, whites today may differ in I.Q. from whites, say, two generations ago as much as whites today differ from blacks today. Given their size and speed, the shifts in time necessarily have been due more to changes in the environment than to changes in the genes. The question then arises: Couldn't the mean of blacks move fifteen points as well through environmental changes? There seems no reason why not—but also no reason to believe that white and Asian means can be made to stand still while the Flynn effect works its magic.

V.

As of 1994, then, we can say nothing for certain about the relative roles that genetics and environment play in the formation of the black-white difference in I.Q. All the evidence remains indirect. The heritability of individual differences in I.Q. does not necessarily mean that ethnic differences are also heritable. But those who think that ethnic differences are readily explained by environmental differences haven't been tough-minded enough about their own argument. At this complex intersection of complex factors, the easy answers are unsatisfactory ones.

Given the weight of the many circumstantial patterns, it seems improbable to us—though possible—that genes have no role whatsoever. What might the mix of genetic and environmental influences be? We are resolutely agnostic on that.

Here is what we hope will be our contribution to the discussion. We put it in italics; if we could, we would put it in neon lights: The answer doesn't much matter. Whether the black-white difference in test scores is produced by genes or the environment has no bearing on any of the reasons why the black-white difference is worth worrying about. If tomorrow we knew beyond a shadow of a doubt what role, if any, were played by genes, the news would be neither good if ethnic differences were predominantly environmental, nor awful if they were predominantly genetic.

The first reason for this assertion is that what matters is not whether differences are environmental or genetic, but how hard they are to change. Many people have a fuzzy impression that if cognitive ability has been depressed by a disadvantaged environment, it is easily remedied. Give the small child a more stimulating environment, give the older child a better education, it is thought, and the environmental deficit can be made up. This impression is wrong. The environment unquestionably has an impact on cognitive ability, but a record of interventions going back more than fifty years has demonstrated how difficult it is to manipulate the environment so that cognitive functioning is improved. The billions of dollars spent annually on compensatory education under Title I of the Elementary and Secondary Education Act have had such a dismal evaluation record that improving general cognitive functioning is no longer even a goal. Preschool education fares little better. Despite extravagant claims that periodically get their fifteen minutes of fame, preschool education, including not just ordinary Head Start but much more intensive programs such as Perry Preschool, raises I.Q. scores by a few points on the exit test, and even those small gains quickly fade. Preschool programs may be good for children in other ways, but they do not have important effects on intelligence. If larger effects are possible, it is only through truly heroic efforts, putting children into full-time, year-round, highly enriched day care from within a few months of birth and keeping them there for the first five years of life—and even

those effects, claimed by the Milwaukee Program and the Abecedarian Project, are subject to widespread skepticism among scholars.

In short: if it were proved tomorrow that ethnic differences in test scores were entirely environmental, there would be no reason to celebrate. That knowledge would not suggest a single educational, preschool, day care or prenatal program that is not already being tried, and would give no reason to believe that tomorrow's effects from such programs will be any more encouraging than those observed to date. Radically improved knowledge about child development and intelligence is required, not just better implementation of what is already known. No breakthroughs are in sight.

The second reason that the concern about genes is overblown is the mistaken idea that genes mean there is nothing to be done. On the contrary, the distributions of genetic traits in a population can change over time, because people who die are not replaced one-for-one by babies with matched dna. Just because there might be a genetic difference among groups in this generation does not mean that it cannot shrink. Nor, for that matter, does genetic equality in this generation mean that genetic differences might not arise within a matter of decades. It depends on which women in which group have how many babies at what ages. More broadly, genetic causes do not leave us helpless. Myops see fine with glasses and many bald men look as if they have hair, however closely myopia and baldness are tied to genes. Check out visual aids and gimmicks on any Macintosh computer to see how technology can compensate for innumeracy and illiteracy.

Now comes the third reason that the concern about genes needs rethinking. It is to us the most compelling: there is no rational reason why any encounter between individuals should be affected in any way by the knowledge that a group difference is genetic instead of environmental. Suppose that the news tomorrow morning is that the black-white difference in cognitive test scores is rooted in genetic differences. Suppose further that tomorrow afternoon, you—let us say you are white—encounter a random African American. Try to think of any way in which anything has changed that should affect your evaluation of or response to that individual and you will soon arrive at a truth that ought to be assimilated by everyone: nothing has changed. That an individual is a member of a group with a certain genetically based mean and distribution in any characteristic, whether it be height, intelligence, predisposition to schizophrenia or eye color has no effect on that reality of that individual. A five-foot man with six-foot parents is still five feet tall, no matter how much height is determined by genes. An African American with an I.Q. of 130 still has an I.Q. of 130, no matter what the black mean may be or to what extent I.Q. is determined by genes. Maybe for some whites, behavior toward black individuals would change if it were known that certain ethnic differences were genetic—but not for any good reason.

We have been too idealistic, one may respond. In the real world, people treat individuals according to their membership in a group. Consider the young black male trying to catch a taxi. It makes no difference how honest he is; many taxi drivers will refuse to pick him up because young black males disproportionately account for taxi robberies. Similarly, some people fear that talking about group differences in I.Q. will encourage employers to use ethnicity as an inexpensive screen if they can get away with it, not bothering to consider black candidates.

These are authentic problems that need to be dealt with. But it puzzles us to hear them raised as a response to the question, "What difference does it make if genes are involved?" Two separate issues are being conflated: the reality of a difference versus its source. An employer has no more incentive to discriminate by ethnicity if he knows that a difference in ability is genetic than if he knows it is "only" environmental. To return to an earlier point,

the key issue is how intractable the difference is. By the time someone is applying for a job, his cognitive functioning can be tweaked only at the margins, if at all, regardless of the original comparative roles of genes and environment in producing that level of cognitive functioning. The existence of a group difference may make a difference in the behavior of individuals toward other individuals, with implications that may well spill over into policy, but the source of the difference is irrelevant to the behavior.

VI.

In *The Bell Curve*, we make all of the above points, document them fully and are prepared to defend them against all comers. We argue that the best and indeed only answer to the problem of group differences is an energetic and uncompromising recommitment to individualism. To judge someone except on his or her own merits was historically thought to be un-American, and we urge that it become so again.

But as we worked on the discussion in the book, we also became aware that ratiocination is not a sufficient response. Many people instinctively believe that genetically caused group differences in intelligence must be psychologically destructive in a way that environmentally caused differences are not. In a way, our informal survey of elites during the writing of the book confirmed this. No matter what we said, we found that people walked away muttering that it does make a difference if genes are involved. But we nonetheless are not persuaded. It seems to us that, on the contrary, human beings have it in them to live comfortably with all kinds of differences, group and individual alike.

We did not put those thoughts into the book. Early on, we decided that the passages on ethnic differences in intelligence had to be inflexibly pinned to data. Speculations were out, and even provocative turns of phrase had to be guarded against. The thoughts we are about to express are decidedly speculative, and hence did not become part of our book. But if you will treat them accordingly, we think they form the basis of a conversation worth beginning, and we will open it here.

As one looks around the world at the huge variety of ethnic groups that have high opinions of themselves, for example, one is struck by how easy it is for each of these clans, as we will call them, to conclude that it has the best combination of genes and culture in the world. In each clan's eyes, its members are blessed to have been born who they are—Arab, Chinese, Jew, Welsh, Russian, Spanish, Zulu, Scots, Hungarian. The list could go on indefinitely, breaking into ever smaller groups (highland Scots, Glaswegians, Scotch-Irish). The members of each clan do not necessarily think their people have gotten the best break regarding their political or economic place in the world, but they do not doubt the intrinsic, unique merits of their particular clan.

How does this clannish self-esteem come about? Any one dimension, including intelligence, clearly plays only a small part. The self-esteem is based on a mix of qualities. These packages of qualities are incomparable across clans. The mixes are too complex, the metrics are too different, the qualities are too numerous to lend themselves to a weighting scheme that everyone could agree upon. The Irish have a way with words; the Irish also give high marks to having a way with words in the pantheon of human abilities. The Russians see themselves as soulful; they give high marks to soulfulness. The Scotch-Irish who moved to America tended to be cantankerous, restless and violent. Well, say the American Scotch-Irish proudly, these qualities made for terrific pioneers.

We offer this hypothesis: Clans tend to order the world, putting themselves on top, not because each clan has an inflated idea of its own virtues, but because each is using a weighting algorithm that genuinely works out that way. One of us had a conversation with a

Thai many years ago about the Thai attitude toward Americans. Americans have technology and capabilities that the Thais do not have, he said, just as the elephant is stronger than a human. "But," he said with a shrug, "who wants to be an elephant?" We do not consider his view quaint. There is an internally consistent logic that legitimately might lead a Thai to conclude that being born Thai gives one a better chance of becoming a complete human being than being born American. He may not be right, but he is not necessarily wrong.

If these observations have merit, why is it that one human clan occasionally develops a deep-seated sense of ethnic inferiority vis-a-vis another clan? History suggests that the reasons tend to be independent of any particular qualities of the two groups, but instead are commonly rooted in historical confrontations. When one clan has been physically subjugated by another, the psychological reactions are complex and long-lasting. The academic literature on political development is filled with studies of the reactions of colonized peoples that prove this case. These self-denigrating reactions are not limited to the common people; if anything, they are most profound among the local elites. Consider, for example, the deeply ambivalent attitudes of Indian elites toward the British. The Indian cultural heritage is glittering, but that heritage was not enough to protect Indian elites from the psychological ravages of being subjugated.

Applying these observations to the American case and to relations between blacks and whites suggests a new way of conceptualizing the familiar "legacy of slavery" arguments. It is not just that slavery surely had lasting effects on black culture, nor even that slavery had a broad negative effect on black self-confidence and self-esteem, but more specifically that the experience of slavery perverted and stunted the evolution of the ethnocentric algorithm that American blacks would have developed in the normal course of events. Whites did everything in their power to explain away or belittle every sign of talent, virtue or superiority among blacks. They had to—if the slaves were superior in qualities that whites themselves valued, where was the moral justification for keeping them enslaved? And so everything that African Americans did well had to be cast in terms that belittled the quality in question. Even to try to document this point leaves one open to charges of condescension, so successfully did whites manage to coopt the value judgments. Most obviously, it is impossible to speak straightforwardly about the dominance of many black athletes without being subject to accusations that one is being backhandedly anti-black.

The nervous concern about racial inferiority in the United States is best seen as a variation on the colonial experience. It is in the process of diminishing as African Americans define for themselves that mix of qualities that makes the American black clan unique and (appropriately in the eyes of the clan) superior. It emerges in fiction by black authors and in a growing body of work by black scholars. It is also happening in the streets. The process is not only normal and healthy; it is essential.

In making these points, there are several things we are not saying that need to be spelled out. We are not giving up on the melting pot. Italians all over America who live in neighborhoods without a single other Italian, and who may technically have more non-Italian than Italian blood, continue to take pride in their Italian heritage in the ways we have described. The same may be said of other ethnic clans. For that matter, we could as easily have used the examples of Texans and Minnesotans as of Thais and Scotch-Irish in describing the ways in which people naturally take pride in their group. Americans often see themselves as members of several clans at the same time—and think of themselves as 100 percent American as well. It is one of America's most glorious qualities.

We are also not trying to tell African Americans or anyone else what qualities should be weighted in their algorithm. Our point is precisely the opposite: no one needs to tell any clan how to come up with a way of seeing itself that is satisfactory; it is one of those things that human communities know how to do quite well when left alone to do it. Still less are we saying that the children from any clan should not, say, study calculus because studying calculus is not part of the clan's heritage. Individuals strike out on their own, making their way in the Great World according to what they bring to their endeavors as individuals—and can still take comfort and pride in their group affiliations. Of course there are complications and tensions in this process. The tighter the clan, the more likely it is to look suspiciously on their children who depart for the Great World—and yet also, the more proudly it is likely to boast of their successes once they have made it, and the more likely that the children will one day restore some of their ties with the clan they left behind. This is one of the classic American dramas.

We are not preaching multiculturalism. Our point is not that everything is relative and the accomplishments of each culture and ethnic group are just as good as those of every other culture and ethnic group. Instead, we are saying a good word for a certain kind of ethnocentrism. Given a chance, each clan will add up its accomplishments using its own weighting system, will encounter the world with confidence in its own worth and, most importantly, will be unconcerned about comparing its accomplishments line-by-line with those of any other clan. This is wise ethnocentrism.

In the context of intelligence and I.Q. scores, we are urging that it is foolish ethnocentrism on the part of European Americans to assume that mean differences in I.Q. among ethnic groups must mean that those who rank lower on that particular dimension are required to be miserable about it—all the more foolish because the group I.Q. of the prototypical American clan, white Protestants, is some rungs from the top.

It is a difficult point to make persuasively, because the undoubted reality of our era is that group differences in intelligence are intensely threatening and feared. One may reasonably ask what point there is in speculating about some better arrangement in which it wouldn't matter. And yet there remain stubborn counterfactuals that give reason for thinking that inequalities in intelligence need not be feared—not just theoretically, but practically.

We put it as a hypothesis that lends itself to empirical test: hardly anyone feels inferior to people who have higher I.Q.s. If you doubt this, put it to yourself. You surely have known many people who are conspicuously smarter than you are, in terms of sheer intellectual horsepower. Certainly we have. There have been occasions when we thought it would be nice to be as smart as these other people. But, like the Thai who asked, "Who wants to be an elephant?" we have not felt inferior to our brilliant friends, nor have we wanted to trade places with them. We have felt a little sorry for some of them, thinking that despite their high intelligence they lacked other qualities that we possessed and that we valued more highly than their extra I.Q. points.

When we have remarked upon this to friends, their reaction has often been, "That's fine for you to say, because you're smart enough already." But we are making a more ambitious argument: it is not just people with high I.Q.s who don't feel inferior to people with even higher I.Q.s. The rule holds true all along the I.Q. continuum.

It is hard to get intellectuals to accept this, because of another phenomenon that we present as a hypothesis, but are fairly confident can be verified: people with high I.Q.s tend to condescend to people with lower I.Q.s. Once again, put yourself to the test. Suppose we point to a person with an I.Q. thirty points lower than yours. Would you be willing to trade places

with him? Do you instinctively feel a little sorry for him? Here, we have found the answers from friends to be more reluctant, and usually a little embarrassed, but generally they have been “no” and “yes,” respectively. Isn’t it remarkable: just about everyone seems to think that his level of intelligence is enough, that any less than his isn’t as good, but that any more than his isn’t such a big deal.

In other words, we propose that the same thing goes on within individuals as within clans. In practice, not just idealistically, people do not judge themselves as human beings by the size of their I.Q.s. Instead, they bring to bear a multidimensional judgment of themselves that lets them take satisfaction in who they are. Surely a person with an I.Q. of 90 sometimes wishes he had an I.Q. of 120, just as a person with an I.Q. of 120 sometimes wishes he had an I.Q. of 150. But it is presumptuous, though a curiously common presumption among intellectuals, to think that someone with an I.Q. of 90 must feel inferior to those who are smarter, just as it is presumptuous to think a white person must feel threatened by a group difference that probably exists between whites and Japanese, a gentile must feel threatened by a group difference that certainly exists between gentiles and Jews or a black person must feel threatened by a group difference between blacks and whites. It is possible to look ahead to a world in which the glorious hodgepodge of inequalities of ethnic groups—genetic and environmental, permanent and temporary—can be not only accepted but celebrated.

This difficult topic calls up an unending sequence of questions. How can intelligence be treated as just one of many qualities when the marketplace puts such a large monetary premium on it? How can one hope that people who are on the lower end of the I.Q. range find places of dignity in the world when the niches they used to hold in society are being devalued? Since the world tends to be run by people who are winners in the I.Q. lottery, how can one hope that societies will be structured so that the lucky ones do not continually run society for their own benefit?

These are all large questions, exceedingly complex questions—but they are no longer about ethnic variations in intelligence. They are about human variation in intelligence. They, not ethnic differences, are worth writing a book about—and that’s what we did. Ethnic differences must be dreaded only to the extent that people insist on dreading them. People certainly are doing so—that much is not in dispute. What we have tried to do here, in a preliminary and no doubt clumsy way, is to begin to talk about the reasons why they need not.