

FROM RICHARD HERRNSTEIN AND CHARLES MURRAY, *THE BELL CURVE: INTELLIGENCE AND CLASS STRUCTURE IN AMERICAN LIFE* (NY: FREE PRESS, 1994).

Chapter 13

Ethnic Differences in Cognitive Ability

Despite the forbidding air that envelops the topic, ethnic differences in cognitive ability are neither surprising nor in doubt. Large human populations differ in many ways, both cultural and biological. It is not surprising that they might differ at least slightly in their cognitive characteristics. That they do is confirmed by the data on ethnic differences in cognitive ability from around the world. One message of this chapter is that such differences are real and have consequences. Another is that the facts are not as alarming as many people seem to fear.

East Asians (e.g., Chinese, Japanese), whether in America or in Asia, typically earn higher scores on intelligence and achievement tests than white Americans. The precise size of their advantage is unclear; estimates range from just a few to ten points. A more certain difference between the races is that East Asians have higher nonverbal intelligence than whites while being equal, or perhaps slightly lower, in verbal intelligence.

The difference in test scores between African-Americans and European-Americans as measured in dozens of reputable studies has converged on approximately a one standard deviation difference for several decades. Translated into centiles, this means that the average white person tests higher than about 84 percent of the population of blacks and that the average black person tests higher than about 16 percent of the population of whites.

The average black and white differ in IQ at every level of socioeconomic status (SES), but they differ more at high levels of SES than at low levels. Attempts to explain the difference in terms of test bias have failed. The tests have approximately equal predictive force for whites and blacks.

In the past few decades, the gap between blacks and whites narrowed by perhaps three IQ points. The narrowing appears to have been mainly caused by a shrinking number of very low scores in the black population rather than

an increasing number of high scores. Improvements in the economic circumstances of blacks, in the quality of the schools they attend, in better public health, and perhaps also diminishing racism may be narrowing the gap.

The debate about whether and how much genes and environment have to do with ethnic differences remains unresolved. The universality of the contrast in nonverbal and verbal skills between East Asians and European whites suggests, without quite proving, genetic roots. Another line of evidence pointing toward a genetic factor in cognitive ethnic differences is that blacks and whites differ most on the tests that are the best measures of *g*, or general intelligence. On the other hand, the scores on even highly *g*-loaded tests can be influenced to some extent by changing environmental factors over the course of a decade or less. Beyond that, some social scientists have challenged the premise that intelligence tests have the same meaning for people who live in different cultural settings or whose forebears had very different histories.

Nothing seems more fearsome to many commentators than the possibility that ethnic and race differences have any genetic component at all. This belief is a fundamental error. Even if the differences between races were entirely genetic (which they surely are not), it should make no practical difference in how individuals deal with each other. The real danger is that the elite wisdom on ethnic differences—that such differences cannot exist—will shift to opposite and equally unjustified extremes. Open and informed discussion is the one certain way to protect society from the dangers of one extreme view or the other.

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. That many readers have turned first to this chapter indicates how sensitive the issue has become.

Our primary purpose is to lay out a set of statements, as precise as the state of knowledge permits, about what is currently known about the size, nature, validity, and persistence of ethnic differences on measures of cognitive ability. A secondary purpose is to try to induce clarity in ways of thinking about ethnic differences, for discussions about such differences tend to run away with themselves, blending issues of fact, theory, ethics, and public policy that need to be separated.

The first thing to remember is that the differences among individu-

als are far greater than the differences between 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. The chapters in Part II, looking only at whites, should have made that clear. But the politics of cognitive inequality get hotter—sometimes too hot to handle—when they are attached to the politics of ethnicity. We believe that the best way to keep the temperature down is to work through the main facts carefully and methodically. This chapter first reviews the evidence bearing on ethnic differences in cognitive ability, then turns to whether the differences originate in genes or in environments. At the chapter's end, we summarize what this knowledge about ethnic differences means in practical terms.

We frequently use the word *ethnic* rather than *race*, because race is such a difficult concept to employ in the American context.^[1] 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 a Latino? A black? The rule we follow here is to classify people according to the way they classify themselves. The studies of “blacks” or “Latinos” or “Asians” who live in America generally denote people who say they are black, Latino, or Asian—no more, no less.

Ethnic Nomenclature

We want to call people whatever they prefer to be called, including their preferences for ethnic labels. As we write, however, there are no hard-and-fast rules. People from Latin America wish to be known according to their national origin: Cuban-American, Mexican-American, Puerto Rican, and so forth. *Hispanic* is still the U.S. government's official label, but *Latino* has gained favor in recent years. We use *Latino*. Opting for common usage and simplicity, we usually use *black* instead of *African-American* and *white* (which always refers to non-Latino whites) instead of *European-American* or *Anglo*. Americans of Asian descent are called *Asian* when the context leaves no possibility of confusion with Asians living in Asia. We shift to the hyphenated versions for everyone when it would avoid such confusions or when, for stylistic reasons, the hyphenated versions seem appropriate.

It would be disingenuous to leave the racial issue at that, however, for race is often on people's minds when they think about IQ. Thus we will eventually comment on cognitive differences among races as they might derive from genetic differences, telling a story that is interesting but still riddled with more questions than answers. This prompts a second point to be understood at the outset: There are differences between races, and they are the rule, not the exception. That assertion may seem controversial to some readers, but it verges on tautology: Races are by definition groups of people who differ in characteristic ways. Intellectual fashion has dictated that all differences must be denied except the absolutely undeniable differences in appearance, but nothing in biology says this should be so. On the contrary, race differences are varied and complex—and they make the human species more adaptable and more interesting.

THE TESTED INTELLIGENCE OF ASIANS, BLACKS, AND WHITES

So much for preliminaries. Answers to commonly asked questions about the ethnic groups in America follow, beginning with the basics and moving into successively more complicated issues. The black-white difference receives by far the most detailed examination because it is the most controversial and has the widest social ramifications. But the most common question we have been asked in recent years has not been about blacks but about Asians, as Americans have watched the spectacular economic success of the Pacific rim nations at a distance and, closer to home, become accustomed to seeing Asian immigrant children collecting top academic honors in America's schools.

Do Asians Have Higher IQs Than Whites?

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. Richard Lynn, a leading scholar of racial and ethnic differences, has reviewed the assembled data on overall Asian IQ in two major articles. In his 1991 review of the literature, he put the median IQ for the studies of Chinese living in Hong Kong, Singapore, Taiwan, and China proper at 110; the median IQ for the studies of Japanese living in Japan at 103; and the median for studies of East Asians living in North America at 103.² But as Lynn acknowledges, these compar-

isons are imprecise because the IQs were not corrected for the changes that have been observed over time in national IQ averages. In Lynn's 1987 compilation, where such corrections were made, the medians for both Chinese and Japanese were 103.³ Mean white American IQ is typically estimated as 101 to 102.⁴ Additional studies of Chinese in Hong Kong, conducted by J. W. C. Chan using the Ravens Standard Progressive Matrices, a nonverbal test that is an especially good measure of *g*, found IQ equivalents in the region of 110 for both elementary and secondary students, compared to about 100 for whites in Hong Kong.⁵ Another study postdating Lynn's review compared representative samples of South Korean and British 9-year-olds and found an IQ difference of nine points.⁶

The most extensive compilation of East Asian cognitive performance in North America, by Philip Vernon, included no attempt to strike an overall estimate for the current gap between the races, but he did draw conclusions about East Asian-white differences in verbal and nonverbal abilities, which we will describe later in the chapter.⁷ In addition to studies of abilities, Vernon compiled extensive data on the schoolwork of East Asians, documenting their superior performance by a variety of measures ranging from grades to the acquisition of the Ph.D. Is this superior performance caused by superior IQ? James Flynn has argued that the real explanation for the success of Asian-Americans is that they are overachievers.⁸ He also says that Asian-Americans actually have the same nonverbal intelligence as whites and a fractionally lower verbal intelligence.⁹ Richard Lynn disagrees and concludes from the same data used by Flynn that there is an ethnic difference in overall IQ as well.¹⁰

The NLSY is not much help on this issue. The sample contained only forty-two East Asians (Chinese, Japanese, and Koreans). Their mean IQ was 106, compared to the European-American white mean of 103, consistent with the evidence that East Asians have a higher IQ than whites but based on such a small sample that not much can be made of it.

The indeterminacy of the debate is predictable. The smaller the IQ difference, the more questionable its reality, and this has proved to be the case with the East Asian-white difference. It is difficult enough to find two sets of subjects within a single city who can be compared without problems of interpretation. Can one compare test scores obtained in different years with different tests for students of different ages in different cultural settings, drawn from possibly different socioeconomic

populations? One answer is that it can be done through techniques that take advantage of patterns observed over many studies. Lynn in particular has responded to each new critique, in some cases providing new data, in others refining earlier estimates, and always pointing to the striking similarity of the results despite the disparity of the tests and settings.¹¹ But given the complexities of crossnational comparisons, the issue must eventually be settled by a sufficient body of data obtained from identical tests administered to populations that are comparable except for race.

We have been able to identify three such efforts. In one, samples of American, British, and Japanese students ages 13 to 15 were administered a test of abstract reasoning and spatial relations. The American and British samples had scores within a point of the standardized mean of 100 on both the abstract and spatial relations components of the test; the Japanese adolescents 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 Vernon for Asians in America.¹²

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—a gap of well over half a standard deviation between both the Japanese and Hong Kong samples and a British one equated for age and socioeconomic status.¹³

The third set of studies, directed by Harold Stevenson, administered 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 Stevenson and his colleagues carefully matched the children on socioeconomic and demographic variables.¹⁴ No significant difference in overall IQ was found, and Stevenson and colleagues 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.”¹⁵

Where does this leave us? The parties in the debate are often individually confident, and you will find in their articles many flat statements that an overall East Asian–white IQ difference does, or does not, exist. We will continue to hedge. Harold Stevenson and his colleagues have convinced us that matching subjects by socioeconomic status can reduce the difference to near zero, but he has not convinced us that

Jews, Latinos, and Gender

In the text we focus on three major racial-ethnic groupings—whites, East Asians, and blacks—because they have dominated both the research and contentions regarding intelligence. But whenever the subject of group differences in IQ comes up, three other questions are sure to be asked: Are Jews really smarter than everyone else? Where do Latinos fit in, compared to whites and blacks? What about women versus men?

Jews—specifically, Ashkenazi Jews of European origins—test higher than any other ethnic group.¹⁶ A fair estimate seems to be that Jews in America and Britain have an overall IQ mean somewhere between a half and a full standard deviation above the mean, with the source of the difference concentrated in the verbal component. In the NLSY, ninety-eight whites with IQ scores identified themselves as Jews. The NLSY did not try to ensure representativeness within ethnic groups other than blacks and Latinos, so we cannot be sure that the ninety-eight Jews in the sample are nationally representative. But it is at least worth noting that their mean IQ was .97 standard deviation above the mean of the rest of the population and .84 standard deviation above the mean of whites who identified themselves as Christian. These test results are matched by analyses of occupational and scientific attainment by Jews, which consistently show their disproportionate level of success, usually by orders of magnitude, in various inventories of scientific and artistic achievement.¹⁷

The term *Latino* embraces people with highly disparate cultural heritages and a wide range of racial stocks. Many of these groups are known to differ markedly in their social and economic profiles. Add to that the problem of possible language difficulties with the tests, and generalizations about IQ become especially imprecise for Latinos. With that in mind, it may be said that their test results generally fall about half to one standard deviation below the national mean. In the NLSY, the disparity with whites was .93 standard deviation. This may be compared to an overall average difference of .84 standard deviation between whites and Mexican-Americans found in the 1960s on the tests used in the famous Coleman report (described in Chapter 17).¹⁸ We will have more to say about the interpretation of Latino scores with regard to possible language bias in Chapter 14.

When it comes to gender, the consistent story has been that men and women have nearly identical mean IQs but that men have a broader distribution. In the NLSY, for example, women had a mean on the Armed Forces Qualification Test (AFQT) that was .06 standard deviation lower than the male mean and a standard deviation that was .11 narrower. For the Wechsler Intelligence Scale for Children, the average boy tests 1.8 IQ points higher than the average girl, and boys have a standard deviation that is .8 point larger than girls.¹⁹ The larger variation among men means that there are more men than women at either extreme of the IQ distribution.

matching by socioeconomic status is a good idea if one wants to know an estimate of the overall difference between East Asians and whites (we will return to the question of matching by socioeconomic status when we discuss comparisons between blacks and whites). In our judgment, the balance of the evidence supports the proposition that the overall East Asian mean is higher than the white mean. If we had to put a number on it, three IQ points currently most resembles a consensus, tentative though it still is. East Asians have a greater advantage than that in a particular kind of nonverbal intelligence, described later in the chapter.

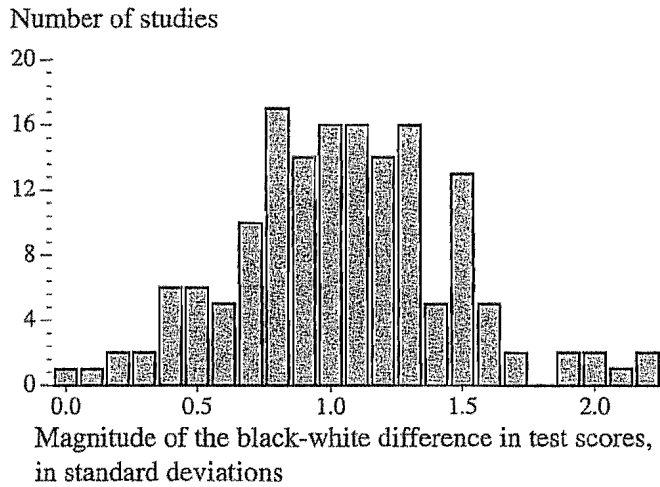
Do Blacks Score Differently from Whites on Standardized Tests of Cognitive Ability?

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 of reliability and validity.²⁰ The answer is also yes for almost all of the studies in which the black and white samples are matched on some special characteristics—samples of juvenile delinquents, for example, or of graduate students—but there are exceptions. The implication of this effect of selecting the groups to be compared is discussed later in the chapter. Since black-white differences are the ones that strain discourse most severely, we will probe deeply into the evidence and its meaning.

How Large Is the Black-White Difference?

The usual answer to this question is one standard deviation.²¹ In discussing IQ tests, for example, the black mean is commonly given as 85, the white mean as 100, and the standard deviation as 15. But the differences observed in any given study seldom conform exactly to one standard deviation. The figure below shows the distribution of the black-white difference (subsequently abbreviated as the “B/W difference”) expressed in standard deviations, in the American studies conducted in this century that have reported the IQ means of a black sample and a white sample and meet basic requirements of interpretability as described in the note.^[22] A total of 156 studies are represented in the plot, and the mean B/W difference is 1.08 standard deviations, or about sixteen IQ points.^[23] The spread of results is substantial, however, reflecting the diversity of the age of the subjects, their geographic loca-

Overview of studies of reporting black-white differences in cognitive test scores, 1918–1990



Sources: Shuey 1966; Osborne and McGurk 1982; Sattler 1988; Vincent 1991; Jensen 1985, 1993b.

tion, their background characteristics, the tests themselves, and sampling error.

When we focus on the studies that meet stricter criteria, the range of values for the B/W difference narrows accordingly. The range of results is considerably reduced, for example, for studies that have taken place since 1940 (after testing's most formative period), outside the South (where the largest B/W differences are found), with subjects older than age 6 (after scores have become more stable), using full test batteries from one of the major IQ tests, and with standard deviations reported for that specific test administration. Of the forty-five studies meeting these criteria, all but nine of the B/W differences are clustered between .5 and 1.5 standard deviations. The mean difference was 1.06 standard deviations, and all but eight of the thirty-one reported a B/W difference greater than .8 standard deviation.

Still 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 1.10 standard deviations. The NLSY, administered in 1980 to by far the largest

sample (6,502 whites, 3,022 blacks) in a national study, found a difference of 1.21 standard deviations on the AFQT.²⁴

Computing the B/W Difference

The simplest way to compute the B/W difference when limited information is available is to take the two means and to compare them using the standard deviation for the reference population, defined in this case as whites. This is how the differences in the figure on page 277 showing the results of 156 studies were computed. When all the data are available, however, as in the case of the NLSY, a more accurate method is available, which takes into account the standard deviations within each population and the relative size of the samples. The equation is given in the note.^[25] Unless otherwise specified, all of the subsequent expressions of the B/W differences are based on this method. (For more about the scoring of IQs in the NLSY, see Appendix 2.)

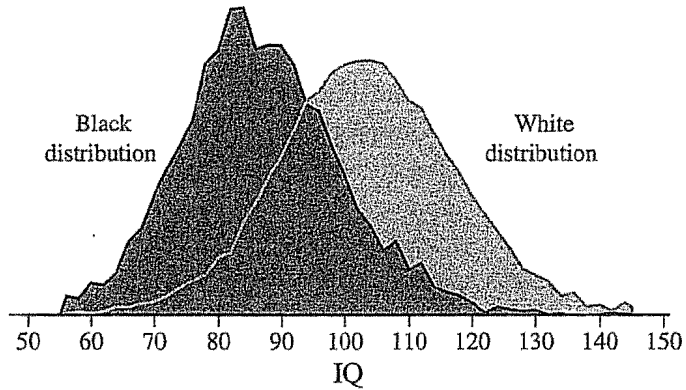
Answering the question "How large is the difference?" in terms of standard deviations does not convey an intuitive sense of the size of the gap. A rough-and-ready way of thinking about the size of the gap is to recall that one standard deviation above and below the mean cuts off the 84th and 16th percentiles of a normal distribution. In the case of the B/W difference of 1.2 standard deviations found in the NLSY, a person with the black mean was at the 11th percentile of the white distribution, and a person with the white mean was at the 91st percentile of the black distribution.

A difference of this magnitude should be thought of in several different ways, each with its own important implications. Recall first that the American black population numbers more than 30 million people. If the results from the NLSY apply to the total black population as of the 1990s, around 100,000 blacks fall into Class I of our five cognitive classes, with IQs of 125 or higher.²⁶ One hundred thousand people is a lot of people. It should be no surprise to see (as one does every day) blacks functioning at high levels in every intellectually challenging field.

It is important to understand as well that 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 below. For any equal number of blacks and whites, a large pro-

The black and white IQ distributions in the NLSY, Version I

Frequency distributions for populations of equal size

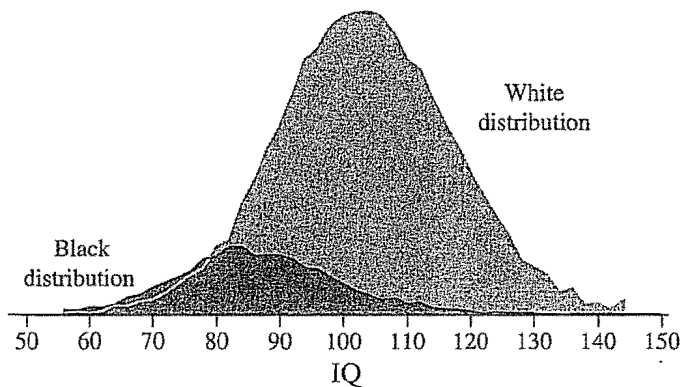


portion have IQs that can be matched up. This is the distribution to keep in mind whenever thinking about individuals.

But an additional complication has to be taken into account: In the United States, there are about six whites for every black. This means that the IQ overlap of the two populations as they actually exist in the United States looks very different from the overlap in the figure just above. The next figure presents the same data from the NLSY when the distributions are shown in proportion to the actual population of young

The black and white IQ distributions in the NLSY, Version II

Frequency distributions proportional to the ethnic composition of the U.S. population



people represented in the NLSY. This figure shows why a B/W difference can be problematic to American society as a whole. At the lower end of the IQ range, there are approximately 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 IQ level are huge. To the extent that the difference represents an authentic difference in cognitive functioning, the social consequences are potentially huge as well. But is the difference authentic?

Are the Differences in Black and White Scores Attributable to Cultural Bias or Other Artifacts of the Test?

Appendix 5 contains a discussion of the state of knowledge regarding test bias. Here, we shall quickly review the basic findings regarding blacks, without repeating the citations in Appendix 5, which we urge you to read.

EXTERNAL EVIDENCE OF BIAS. Tests are used to predict things—most commonly, to predict performance in school or on the job. Chapter 3 discussed this issue in detail. You will recall that 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 kind of bias would be to compensate the black applicants by, for example, adding the appropriate number of points onto their scores.

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 secondary school, in the university, in the armed forces, in unskilled and skilled jobs, in the professions. Overwhelmingly, the evidence is that the major standardized tests used to help make school and job decisions^[27] do not underpredict black performance, nor does the expert community find any other general or systematic difference in the predictive accuracy of tests for blacks and whites.^[28]

INTERNAL EVIDENCE OF BIAS. Predictive validity is the ultimate criterion for bias, because it involves the proof of the pudding for any test. But although predictive validity is in a technical sense the decisive issue, our impression from talking about this issue with colleagues and friends is that other types of potential bias loom larger in their imaginations: the many things that are put under the umbrella label of “cultural 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 answer 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.”^[29] In the language of psychometrics, this is called *internal*

evidence of bias, as contrasted with the external evidence of differential prediction.

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 B/W 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?³⁰ Are differences in group test scores concentrated among certain items?

The technical literature is again clear. In study after study of the leading tests, the hypothesis that the B/W difference is caused by questions with cultural content has been contradicted by the facts.³¹ 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. For groups that have special language considerations—Latinos and American Indians, for example—some internal evidence of bias has been found, unless English is their native language.³²

Studies comparing blacks and whites on various kinds of IQ tests find that the B/W difference is not created by items that ask about regattas or who wrote *Hamlet*, or any of the other similar examples cited in criticisms of tests. 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 B/W difference is 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 both so well established empirically yet comes as such a surprise to most people who are new to this topic. We will elaborate on this finding later in the chapter. In any case, there is no longer an important technical debate over the conclusion that the cultural content of test items is not the cause of group differences in scores.

"MOTIVATION TO TRY." Suppose that 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 is less attuned to testing situations (from one point of view), or is inclined to put up with such nonsense (from another). Perhaps he 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 out his best effort because of the peer pressures against "acting white" in some inner-city schools.

The studies that have attempted to measure motivation in such situations have generally found that blacks are at least as motivated as whites.³³ But these are not wholly convincing, for why should 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. Two sets of data seem especially pertinent.

The first involves 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 in concept, uses numbers that are familiar to everyone, and calls on no cultural information beyond knowing numbers. The digit span is especially informative regarding motivation not just because of the low cultural loading of the item, but because the backward form is twice as *g*-loaded as the forward form, a much better measure of general intelligence. The reason is that repeating the numbers is mentally more demanding than repeating them in the heard order, as readers can determine for themselves by a little self-testing.

The two parts of the subtest have identical content. They are administered at the same time during the test. Each subject does both. But in studies the black-white difference is about twice as great on backward digits as on forward digits.³⁴ The question arises: How can lack of motivation (or test willingness or any other explanation of that kind) explain the difference in performance on the two parts of the same test?³⁵

A similar question arises from work on reaction time. Several psychometricians, led by Arthur Jensen, have been exploring the underlying nature of *g* by hypothesizing that neurologic processing speed is implicated, akin to the speed of the microprocessor in a computer.

Smarter people process faster than less smart people. The strategy for testing the hypothesis is to give people extremely simple cognitive tasks—so simple that no conscious thought is involved—and to use precise timing methods to determine how fast different people perform these simple tasks. One commonly used apparatus involves a console with a semicircle of eight lights, each with a button next to it. In the middle of the console is the “home” button. At the beginning of each trial, the subject is depressing the home button with his finger. One of the lights in the semicircle goes on. The subject moves his finger to the button closest to the light, which turns it off. There are more complicated versions of the task (three lights go on, and the subject moves to the one that is farthest from the other two, for example), but none requires much thought, and everybody gets every trial “right.” The subject’s response speed is broken into two measurements: reaction time (RT), the time it takes the subject to lift his finger from the home button after a target light goes on, and movement time (MT), the time it takes to move the finger from just above the home button to the target button.³⁶

Francis Galton in the nineteenth century believed that reaction time is associated with intelligence but could not prove it. He was on the right track after all. In modern studies, reaction time is correlated with the results from full-scale IQ tests; even more specifically, it is correlated with the *g* factor in IQ tests—in some studies, *only* with the *g* factor.³⁷ Movement time is much less correlated with IQ or with *g*.³⁸ This makes sense: Most of the cognitive processing has been completed by the time the finger leaves the home button; the rest is mostly a function of small motor skills.

Research on reaction time is doing much to advance our understanding of the biological basis of *g*. For our purposes here, however, it also offers a test of the motivation hypothesis: The consistent result of many studies is that white reaction time is faster than black reaction time, but black movement time is faster than white movement time.³⁹ One can imagine an unmotivated subject who thinks the reaction time test is a waste of time and does not try very hard. But the level of motivation, whatever it may be, seems likely to be the same for the measures of RT and MT. The question arises: How can one be unmotivated to do well during one split-second of a test but apparently motivated during the next split-second? Results of this sort argue against easy explanations that appeal to differences in motivation as explanatory of the B/W difference.

UNIFORM BACKGROUND BIAS. Other kinds of bias discussed in Appendix 5 include the possibility that blacks have less access to coaching than whites, less experience with tests (less "testwiseness"), poorer understanding of standard English, and that their performance is affected by white examiners. Each of these hypotheses has been investigated, for many tests, under many conditions. None has been sustained. In short, the testable hypotheses have led toward the conclusion that cognitive ability tests are not biased against blacks. This leaves one final hypothesis regarding cultural bias that does not lend itself to empirical evaluation, at least not directly.

Suppose our society is so steeped in the conditions that produce test bias that people in disadvantaged groups underscore their cognitive abilities on *all* the items on tests, thereby hiding the internal evidence of bias. At the same time and for the same reasons, they underperform in school and on the job in relation to their true abilities, thereby hiding the external evidence. In other words, the tests may be biased against disadvantaged groups, but the traces of bias are invisible because the bias permeates all areas of the group's performance. Accordingly, it would be as useless to look for evidence of test bias as it would be for Einstein's imaginary person traveling near the speed of light to try to determine whether time has slowed. Einstein's traveler has no clock that exists independent of his space-time context. In assessing test bias, we would have no test or criterion measure that exists independent of this culture and its history. This form of bias would pervade everything.

To some readers, the hypothesis will seem so plausible that it is self-evidently correct. Before deciding that this must be the explanation for group differences in test scores, however, a few problems must be overcome. First, the comments about the digit span and reaction time results apply here as well. How can this uniform background bias suppress black reaction time but not the movement time? How can it suppress performance on backward digit span more than forward digit span? Second, the hypothesis implies that many of the performance yardsticks in the society at large are not only biased, they are all so similar in the degree to which they distort the truth—in every occupation, every type of educational institution, every achievement measure, every performance measure—that no differential distortion is picked up by the data. Is this plausible?

It is not good enough to accept without question that a general "background radiation" of bias, uniform and ubiquitous, explains away black

and white differences in test scores and performance measures. The hypothesis might, in theory, be true. But given the degree to which everyday experience suggests that the environment confronting blacks in different sectors of American life is not uniformly hostile and given the consistency in results from a wide variety of cognitive measures, assuming that the hypothesis is true represents a considerably longer leap of faith than the much more limited assumption that race prejudice is still a factor in American life. In the matter of test bias, this brings us to the frontier of knowledge.

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. We will take up the two answers and their associated rationales separately:

First version: *If you extract the effects of socioeconomic class, what happens to the overall magnitude of the B/W difference?* Blacks are disproportionately in the lower socioeconomic classes, and socioeconomic class is known to be associated with IQ. Therefore, many people suggest, part of what appears to be an ethnic difference in IQ 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 B/W difference in the NLSY is 1.21. In a regression equation in which both race and socioeconomic background are entered, the difference between whites and blacks shrinks to .76 standard deviation.^[40] Socioeconomic status explains 37 percent of the original B/W difference. This relationship is in line with the results from many other studies.^[41]

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.^[42] The trouble is that socioeconomic status is also a *result* of cognitive ability, as people of high and low cognitive ability move to correspondingly high and low places in the socioeconomic continuum. The reason that parents have high or low socioeconomic status is in part

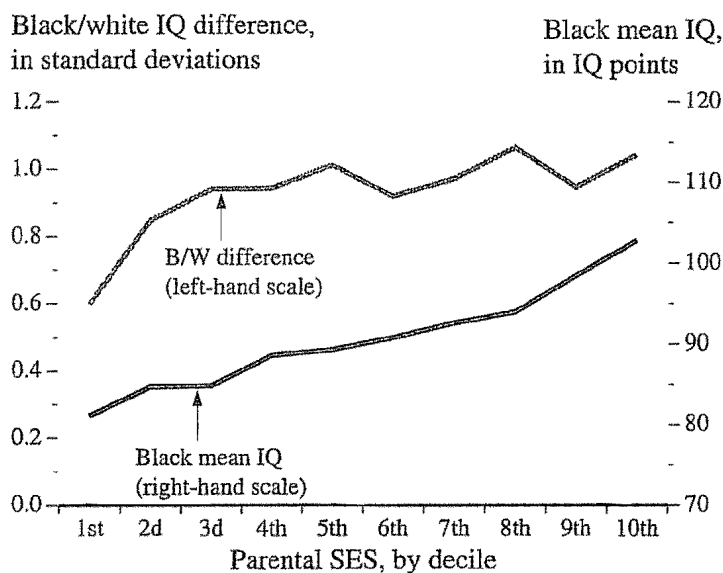
a function of their intelligence, and their intelligence also affects the IQ of the children via both genes and environment.

Because of these relationships, "controlling" for socioeconomic status in racial comparisons is guaranteed to reduce IQ differences in the same way that choosing black and white samples from a school for the intellectually gifted is guaranteed to reduce IQ differences (assuming race-blind admissions standards). But the remaining difference is not necessarily more real or authentic than the one we start with. This seems to be a hard point to grasp, judging from the pervasiveness of controlling for socioeconomic status in the sociological literature on ethnic differences. But suppose we were asking whether blacks and whites differed in sprinting speed, and controlled for "varsity status" by examining only athletes on the track teams in Division I colleges. Blacks would probably still sprint faster than whites on the average, but it would be a smaller difference than in the population at large. Is there any sense in which this smaller difference would be a more accurate measure of the racial difference in sprinting ability than the larger difference in the general population? We pose that as an interesting theoretical issue. In terms of numbers, a reasonable rule of thumb is that controlling for socioeconomic status reduces the overall B/W difference by about a 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/IQ question referred to the overall score of a population of blacks and whites. The second version concentrates on the B/W difference within socioeconomic classes. The rationale goes like this: Blacks score lower on average because they are socioeconomically at a disadvantage in our society. This disadvantage should most seriously handicap the children of blacks in the lower socioeconomic classes, who suffer from greater barriers to education and occupational advancement than do the children of blacks in the middle and upper classes. As blacks advance up the socioeconomic ladder, their children, less exposed to these environmental deficits, 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 by using our parental SES index and matching it against the mean IQ score, as shown in the figure below. IQ scores increase with economic status for both races. But as the figure shows, the magnitude of the B/W difference in standard deviations does not decrease. Indeed, it gets larger as people move up from the very bottom of the socioeco-

Black IQ scores go up with socioeconomic status, but the black-white difference does not shrink



conomic 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.^[43]

How Do African-Americans Compare with Blacks in Africa on Cognitive Tests?

This question often arises in the context of black-white comparisons in America, the thought being that the African black population has not been subjected to the historical legacy of American black slavery and discrimination and might therefore have higher scores. Many studies of African students in primary and secondary schools, in both urban and rural areas, have included cognitive ability tests. As in the United States, it has been demonstrated in Africa that the same test items that discriminate best among blacks discriminate best among whites and that the same factors that depress white scores (for example, coming from a rural area) depress black scores. The predictive validity of tests for academic and job performance seems to be about the same. In general, the

psychometric properties of the standardized tests are the same for blacks living in Africa as for American blacks.⁴⁴

It has been more difficult to assemble data on the score of the average African black than one would expect, given the extensiveness of the test experience in Africa. In the same review of the literature that permitted the above generalizations, for example—a thirty-page article followed by a bibliography of more than 200 titles—not a single average is reported.⁴⁵ One reason for this reluctance to discuss averages is that blacks in Africa, including urbanized blacks with secondary educations, have obtained extremely low scores. Richard Lynn was able to assemble eleven studies in his 1991 review of the literature. He estimated the median black African IQ to be 75, approximately 1.7 standard deviations below the U.S. overall population average, about ten points lower than the current figure for American blacks.⁴⁶ Where other data are available, the estimates of the black African IQ fall at least that low and, in some instances, even lower.⁴⁷ The IQ of “coloured” students in South Africa—of mixed racial background—has been found to be similar to that of American blacks.⁴⁸

In summary: African blacks are, on average, substantially below African-Americans in intelligence test scores. Psychometrically, there is little reason to think that these results mean anything different about cognitive functioning than they mean in non-African populations. For our purposes, the main point is that the hypothesis about the special circumstances of American blacks depressing their test scores is not substantiated by the African data.

Is the Difference in Black and White Test Scores Diminishing?

The answer is yes with (as usual) some qualifications.

IQ TEST DATA. The most straightforward way to answer the question would be to examine the repeated administrations of the same IQ tests to comparable populations, but large, nationally representative IQ data are not produced every year (or even every decade). The NLSY data are among the most recent for a young adult population, and they have a B/W difference toward the high end of the range. The only post-1980 study reporting black and white adult averages that we have found is the renorming of the Wechsler Adult Intelligence Scale (WAIS-R) in 1981 in which the difference between blacks and a sample of whites

(that apparently did not try to discriminate between Latino and Anglo whites) was 1.0 standard deviation.⁴⁹

Recent data on children tell opposite stories. In a review of IQ tests of children conducted since 1980, Ken Vincent of the University of Houston reports results for four normative studies that showed a B/W difference of only seven IQ points for the Ravens Standard Progressive Matrices (SPM) and the Kaufman Assessment Battery for Children (K-ABC).⁵⁰ Two other studies involving the Stanford-Binet IV found B/W differences of ten points for children ages 7 to 11 and twelve points for children ages 2 to 6.^[51] Qualifications must be attached to these findings. The B/W difference on the K-ABC normative sample has in particular been subjected to reexamination suggesting that the diminished gap largely reflected psychometric and statistical artifacts.^[52] Nonetheless, the data on children that Vincent reviews may be read as encouraging. The most impressive of the findings is the comparatively small B/W difference of only seven IQ points on the Ravens SPM administered to 12-year-olds. This finding corresponds to Jensen's 1992 study of black and white children in an upper-middle-class setting in which the difference on the Ravens SPM was similarly below the norm (a deficit corresponding to ten IQ points).⁵³

In contrast to Vincent's optimistic conclusions, the NLSY shows a growing rather than a shrinking gap in the next generation of blacks and whites. As discussed in Chapter 15, the B/W difference between NLSY children is currently wider than the B/W difference separating their mothers.

ACADEMIC APTITUDE AND ACHIEVEMENT TESTS. The most extensive evidence of a narrowing black-white gap can be found in longitudinal data from the National Assessment of Educational Progress (NAEP), the American College Testing (ACT) examination, the SAT, a comparison of the 1972 and 1980 national high school surveys, and some state-level achievement test data. We review the NAEP and the SAT here, and others (which tell the same story) in Appendix 5.

The National Assessment of Educational Progress is an ongoing program sponsored by the federal government to monitor the academic achievement of the nation's youth. It began in 1969, periodically testing 9-, 13-, and 17-year-olds in science, mathematics, reading, and writing in nationally representative samples. The table below shows the changes from the first round of testing in 1969-1973 to the data for

**Reductions in the Black-White Difference on the
National Assessment of Educational Progress**

	White-Black Difference, in Standard Deviations ^a		Change
	1969-1973	1990	
<i>9-year-olds</i>			
Science	1.14	.84	-.30
Math	.70	.54	-.16
Reading	.88	.70	-.18
<i>Average</i>	.91	.69	-.21
<i>13-year-olds</i>			
Science	.96	.76	-.20
Math	.92	.54	-.38
Reading	.78	.40	-.38
<i>Average</i>	.89	.57	-.32
<i>17-year-olds</i>			
Science	1.08	.96	-.12
Math	.80	.42	-.38
Reading	1.04	.60	-.44
<i>Average</i>	.97	.66	-.31
<i>Overall average</i>	.92	.64	-.28

Source: National Center for Education Statistics, 1991b.

^a The computations assume a standard deviation of 50.

1990, expressed in standard deviations. The "Change" column gives the earlier B/W difference minus the later B/W difference, which is negative if the gap is closing. The fourth component of the NAEP, a writing test, was introduced only in 1984, with replications in 1988 and 1990. Unlike all the others, it does not show a narrowing of the white-black gap (.46 SD in both 1984 and 1990) and is not included in the table.

As the table indicates, black progress in narrowing the test score discrepancy with whites has been substantial on all three tests and across all of the age groups. The overall average gap of .92 standard deviation in the 1969-1973 tests had shrunk to .64 standard deviation by 1990. The gap narrowed because black scores rose, not because white scores fell. Altogether, the NAEP provides an encouraging picture.

The first published breakdowns of SAT scores by ethnicity appear for 1976, when the downward trend in SAT scores nationwide after 1963 was nearing its bottom (see Chapter 18). From 1976 to 1993, the white-black gap in SAT scores narrowed from 1.16 to .88 standard deviation in the verbal portion of the test and from 1.27 to .92 standard deviation in the mathematics portion of the test.^[54] Comparable narrowing has also brought black and white achievement test scores closer, as presented in Appendix 5. Because the ethnic self-identification of SAT test takers contains some anomalies⁵⁵ and because the SAT pool is unrepresentative of the general population, the numbers should be interpreted with caution. But even so, the SAT data indicate a narrowing gap. Black SAT test takers improved substantially more in scores than white SAT test takers, and neither the changes in the pool of test takers nor the well-advertised national decline in SAT scores was responsible, for reasons explained in the notes.^[56]

EXPLAINING THE CONVERGENCE. Let us assume that during the past two decades black and white cognitive ability as measured by IQ has in fact converged by an amount that is consistent with the convergence in educational aptitude measures such as the SAT or NAEP—a narrowing of approximately .15 to .25 standard deviation units, or the equivalent of two to three IQ points overall.^[57] Why have the scores converged? The answer calls for speculation.

We take for granted that individual variations in cognitive ability depend on both genes and environment (see Chapter 4). In a period as short as twenty years, environmental changes are likely to provide the main reason for the narrowing racial gap in scores.^[58] Real and important though the problems of the underclass are, and acknowledging that the underclass is disproportionately black, living conditions have improved for most African-Americans since the 1950s—socially, economically, and educationally.

Consider the schools that blacks attend, for example. Some schools in the inner cities are worse than they were thirty years ago, but proportionately few blacks live in these worst-of-the-worst areas.⁵⁹ Throughout the South and in much of the rest of the country, many black children as recently as the 1950s attended ramshackle schools with undertrained teachers and meager teaching materials. Any comparison between the schools that most blacks attend now and the ones they attended in the 1950s favors contemporary schools. Assuming that

education affects cognitive capacity, the rising investment in education disproportionately benefits the cognitive levels at the lower end of the socioeconomic spectrum.

The argument can be repeated for public health. If nutrition, shelter, and health care affect intellectual development, then rising standards of living are disproportionately going to show up in rising scores for the economically disadvantaged rather than for the upper classes. For travel and its educational benefits, the argument also applies. Not so long ago, many less advantaged people spent their lives within a few miles of their birthplaces. Today, Americans of nearly all walks of life crowd the interstate roads and the airports. Finally, for that most contemporary form of vicarious travel—the popular media—the leveling is still more dramatic. The modern media can bring the world to everyone in ways that were once open only to the rich.

Because blacks are shifted toward the lower end of the socioeconomic range, such improvements benefit them, on average, more than whites. If the improvements affect cognitive development, the black-white gap should have contracted. Beyond this socioeconomic leveling, there might also have been a leveling due to diminishing racism. The legacy of historic racism may still be taking its toll on cognitive development, but we must allow the possibility that it has lessened, at least for new generations. This too might account for some narrowing of the black-white gap.

LOOKING TO THE FUTURE. The question that remains is whether black and white test scores will continue to converge. If all that separates blacks from whites are environmental differences and if fertility patterns for different socioeconomic groups are comparable, there is no reason why they shouldn't. The process would be very slow, however. If it continues at the pace observed over the last twenty years, then we could expect black and white SAT scores to reach equality sometime in the middle of the twenty-first century, but linear extrapolations over such long periods are not worth much.^[60]

If black fertility is loaded more heavily than white fertility toward low-IQ segments of the population, then at some point convergence may be expected to stop, and the gap could even begin to widen again. We take up the fertility issue in Chapter 15. A brief summary statement concerning fertility patterns is that the news is not good. For now, the test score data leave open the possibility that convergence has already

stalled. For most of the tests we mentioned, black scores stopped rising in the mid-1980s. On the NAEP, the B/W gap actually increased from 1986 to 1990 in all but one test group (the math test for 17-year-olds). On the SAT, black scores on both verbal and math parts were nearly flat for the five years ending in 1993, after substantial gains in the preceding decade. On the ACT, however, black scores continued to rise after 1986, albeit modestly.^[61]

One explanation for the stalled convergence on the NAEP and SAT is that American education stopped improving for everyone, blacks included. This is consistent with the white experience on the SAT, where white scores have also been nearly flat since the mid-1980s. But the logic is suspect. Just because a group at a higher mean stops improving does not imply that a group with a lower mean should also stop improving. On the contrary, pessimists can develop a case that the convergence of black and white SAT scores in the last two decades is symptomatic of what happens when education slows down toward the speed of the slowest ship in the convoy. It may well be that education improves for students at the low end of the distribution but gets worse (or, more optimistically, improves less) for students at the top end.^[62] If that is the case, the gap between people at the low and high end of the distribution should narrow, but the narrowing will stop once the educational system completes its readjustment favoring less capable students.

The narrowing black-white gap on the SAT looks consistent with some such explanation.^[63] Seen from one perspective, there is good news all along the spectrum of test scores. From 1980 to 1993, the proportion of black test takers who scored in the 700s on the SAT-Verbal increased by 27 percent, for example.^[64] But such changes at the high end of the range of test scores mean little, because so small a proportion of all black students were involved.^[65] The real source of the black increase of twenty-three points in the average verbal test score from 1980 to 1993 was a rise in the scores at the low end of the range. More than half (51 percent) of the gain occurred because the proportion of black students scoring in the 200s dropped from 42 percent to 30 percent.^[66] In contrast, less than 1 percent (0.4 percent) of the gain occurred because of the change in the proportion of black students scoring in the 700s. For the math test, 22 percent of the gain from 1980 to 1993 was accounted for by a drop in students scoring in the 200s; 4 percent of it was accounted for by an increase in students scoring in the 700s.

Pessimists reading these data may think of an analogy with the in-

creases in height that follow from better nutrition: Better nutrition helps raise the height of children whose diets would otherwise have been inadequate, but it does not add anything to the height of those who have been receiving a good diet already.^[67] Optimists may use the opposite sort of nutritional analogy: the experience of trying to lose weight. Even a successful diet has its plateaus, when the weight stubbornly stops coming off for a while. A plateau is all that we are seeing in recent test data. Perhaps convergence will resume or even accelerate in the near future.

At the least, the optimists may say that it is too soon to pass judgment, and that seems the safest conclusion. As we reach the end of this discussion of convergence, we can imagine the responses of readers of varying persuasions. Many of you will be wondering why we have felt it necessary to qualify the good news. A smaller number of readers who specialize in mental testing may be wondering why we have given so much prominence to educational achievement trends and a scattering of IQ results that may be psychometrically ephemeral. The answer for everyone is that predicting the future on this issue is little more than guesswork at this point. We urge upon our readers a similar suspension of judgment.

GENETICS, IQ, AND RACE

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 and Stanley Rothman, a psychologist and a political scientist, respectively, sent a questionnaire to a broad sample of 1,020 scholars, mostly academicians, whose specialties give them reason to be knowledgeable about IQ.^[68] Among the other questions, they asked, "Which of the following best characterizes your *opinion* of the heritability of the black-white difference in IQ?" (emphasis in the questionnaire item). 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.

The responses reveal the degree of uncertainty within the scientific community about where the truth lies. We have considered leaving the genetics issue at that, on grounds that no useful purpose is served by talking about a subject that is so inflammatory, so painful, and so far from resolution. We could have cited any number of expert reassurances that genetic differences among ethnic groups are not worth worrying about. For example, a recently published textbook from which college students around the country are learning about intelligence states unequivocally that "there is no convincing direct or indirect evidence in favor of a genetic hypothesis of racial differences in IQ."⁶⁹ Stephen J. Gould, whose *Mismeasure of Man* so successfully cemented the received wisdom about IQ in the media, expresses this view as confidently and more eloquently. "Equality [of the races] is not given a priori," he once wrote in his column for *Natural History* magazine. "It is neither an ethical principle (though equal treatment may be) nor a statement about norms of social action. It just worked out that way. A hundred different and plausible scenarios for human history would have yielded other results (and moral dilemmas) of enormous magnitude. They just didn't happen."⁷⁰ He goes on to make three arguments. First, the very concept of race is illegitimate, given the extensiveness of interbreeding and the imprecise nature of most of the traits that people think of as being "racial." Second, the division of races is recent, occurring only in the last tens or perhaps hundreds of thousands of years, limiting the amount of time that groups of humans could have taken separate evolutionary paths. Third, developments in genetics demonstrate that the genetic differences among human beings are minor. "We now know that our usual metaphor of superficiality—skin deep—is literally accurate," Gould writes.⁷¹ He concludes: "Say it five times before breakfast tomorrow; more important, understand it as the center of a network of implication: 'Human equality [i.e., equality among the races] is a contingent fact of history.'"⁷²

Our difficulty with this position is not that Gould (or others who make similar arguments) is wrong about the blurred lines between the races, or about how long the races have been separated, or about the number of genes that are racially distinctive. All his facts can be true, and yet people who call themselves Japanese or Xhosa or Caucasians or

Maori can still differ intellectually for genetic reasons. We may call them "ethnic groups" instead of races if we wish—we too are more comfortable with *ethnic*, because of the blurred lines—but some ethnic groups nonetheless differ genetically for sure, otherwise they would not have differing skin colors or hair textures or muscle mass. They also differ intellectually on the average. The question remaining is whether the intellectual differences overlap the genetic differences to any extent.

Our reason for confronting the issue of genetic cognitive differences is not to quarrel with those who deny them. If the question of genetic differences in cognitive ability were something that only professors argued about among themselves, we would happily ignore it here. We cannot do so, first because in the public discussion of genes and intelligence, no burden of proof at all is placed on the innumerable public commentators who claim that racial differences in intelligence are purely environmental. This sometimes leads to a next statement: that the differences are therefore inauthentic and that public policy must be measured against the assumption that there are no genuine cognitive differences between the races.⁷³ The assumption of genetic cognitive equality among the races has practical consequences that require us to confront the assumption directly.

Second, we have become convinced that the topic of genes, intelligence, and race in the late twentieth century is like the topic of sex in Victorian England. Publicly, there seems to be nothing to talk about. Privately, people are fascinated by it. As the gulf widens between public discussion and private opinion, confusion and error flourish. As it was true of sex then, so it is true of ethnic differences in intelligence now: Taboos breed not only ignorance but misinformation.

The dangers of the misinformation are compounded by the nature of the contemporary discussion of race. Just beneath the surface of American life, people talk about race in ways that bear little resemblance to the politically correct public discussion. Conducted in the workplace, dorm rooms, taverns, and country clubs, by people in every ethnic group, this dialogue is troubled and often accusatory. The underground conversation is not limited to a racist minority. It goes on everywhere, and we believe is increasingly shaped by privately held beliefs about the implications of genetic differences that could not stand open inspection.

The evidence about ethnic differences can be misused, as many people say to us. Some readers may feel that this danger places a moral pro-

hibition against examining the evidence for genetic factors in public. We disagree, in part because we see even greater dangers in the current gulf between public pronouncements and private beliefs. And so, for better or worse, here are the major strands of current thinking about the role of genes in cognitive differences between races.^[74]

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. As we discussed in Chapter 4, scholars accept that IQ is substantially heritable, somewhere between 40 and 80 percent, meaning that much of the observed variation in IQ is genetic. And yet this information tells us nothing for sure about the origin of the differences between races in measured intelligence. This point is so basic, and so commonly misunderstood, that it deserves emphasis: *That a trait is genetically transmitted in individuals 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.⁷⁵ 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 IQ points) separates American blacks and whites and that a fifth of a standard deviation (three IQ points) separates East Asians and whites. Finally, we assume that IQ 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 IQ 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.¹⁷⁶ Let's state these conclusions in percentile terms: The *average* environment of blacks would have to be at the 6th percentile of the distribution of environments among whites, and the *average* environment of East Asians would have to be at the 63rd percentile of environments among whites, for the racial differences to be entirely environmental.

Environmental differences of this magnitude and pattern are implausible. Recall further that the B/W difference (in standardized units) is smallest at the lowest socioeconomic levels. Why, if the B/W 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. 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.

Environmental explanations may successfully circumvent these problems, but the explanations have to be formulated rather than simply assumed. Our initial objective is to warn readers who come to the discussion with firmly held opinions on either side. The heritability of individual differences in IQ 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.

Reasons for Thinking that Genetic Differences Might Be Involved

Now we turn to some of the more technical arguments, beginning with those that argue for some genetic component in group differences.

PROFILE DIFFERENCES BETWEEN WHITES AND EAST ASIANS. Races differ not just in average scores but in the profile of intellectual capacities. A full-scale IQ score is the aggregate of many subtests. There are thirteen of them in the Wechsler Intelligence Scale for Children (WISC-R), for example. The most basic division of the subtests is into a verbal IQ and a performance IQ. In white samples,

the verbal and performance IQ subscores tend to have about the same mean, because IQ tests have been standardized on predominantly white populations. But individuals can have imbalances between these two IQs. 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 IQs—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 IQ and substantially higher on visuospatial IQ. Even in the rare studies that have found overall Japanese or Chinese IQs no higher than white IQs (e.g., the Stevenson study of Japanese, Taiwanese, and Minnesotans mentioned earlier),⁷⁷ the discrepancy between verbal and visuospatial IQ 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.⁷⁸ A study of a small sample of Korean infants adopted into white families in Belgium found the familiar elevated visuospatial scores.⁷⁹

This finding has an echo in the United States, where Asian-American students abound in engineering, in medical schools, and in graduate programs in the sciences, but are scarce in law schools and graduate programs in the humanities and social sciences. Most people reflexively assume that this can be explained by language differences. People who did not speak English as their first language or who grew up in households where English was not the language of choice choose professions that are not so dependent on fluent English, we often hear. But the explanation becomes less credible with every passing year. Philip Vernon, after reviewing the evidence on Asian-Americans, concluded that unfamiliarity with the English language and American culture is a plausible explanation only for the results of the early studies. Contemporary studies of Asian-Americans who are thoroughly acculturated also show the typical discrepancy in verbal and visuospatial abilities. American Indians and Inuit similarly score higher visuospatially than verbally;

their ancestors migrated to the Americas from East Asia hundreds of centuries ago.⁸⁰ The verbal-visuospatial discrepancy goes deeper than linguistic background.

Vernon's overall appraisal was that the mean Asian-American IQ is about 97 on verbal tests and about 110 on visuospatial tests.^[81] Lynn's 1987 review of the IQ literature on East Asians found a median verbal IQ of 98 and a median visuospatial IQ of 106.^[82] As of 1993, for Asian-American students who reported that English was the first language they learned (alone or with another language), the Asian-American SAT mean was .21 standard deviation above the national mean on the verbal test and .43 standard deviation above the national mean on the math test. Converted to an IQ metric, this amounts to a 3.3 point elevation of mathematical scores over verbal scores for the high IQ Asian-American population that takes the SAT.^[83]

Why do visuospatial abilities develop more than verbal abilities in people of East Asian ancestry in Japan, Hong Kong, Taiwan, mainland China, and other Asian countries and in the United States and elsewhere, despite the differences among the cultures and languages in all those countries? Any simple socioeconomic, cultural, or linguistic explanation is out of the question, given the diversity of living conditions, native languages, educational resources, and cultural practices experienced by Hong Kong Chinese, Japanese in Japan or the United States, Koreans in Korea or Belgium, and Inuit or American Indians. We are not so rash as to assert that the environment or the culture is *wholly* irrelevant to the development of verbal and visuospatial abilities, but the common genetic history of racial East Asians and their North American or European descendants on the one hand, and the racial Europeans and their North American descendants, on the other, cannot plausibly be dismissed as irrelevant.

PROFILE DIFFERENCES BETWEEN WHITES AND BLACKS. Turning now to blacks and whites (using these terms to refer exclusively to Americans), ability profiles have also 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*.^[84] This hypothesis says that if the B/W difference on test scores reflects a real underlying difference in the general mental ability, *g*, then the size of the B/W difference will be related to the degree to which the test is saturated with *g*.^[85] In other words, the better a test measures *g*, the larger the black-

white difference will be. Arthur Jensen began to explore this possibility when he looked at the pattern of subtest scores on the WISC-R, taking advantage of the fact that the WISC-R has thirteen subtests, each measuring a somewhat different skill. Converting their statistical procedures into a more easily understood form, here is the logic of what Arthur Jensen and his coauthor, Cyril Reynolds, did.⁸⁶

On average, low-SES whites get lower test scores than high-SES whites. But suppose you were to go through a large set of white test scores from a low-SES and a high-SES group and pull out everyone with an overall IQ score of, say, 105. Now you have identical scores but very different SES groups. The question becomes, What does the pattern of subtest scores look like? The answer is, The same. Once you equalize the overall IQ scores, low-SES and high-SES whites also had close-to-identical mean scores on the individual subtests.

Now do the same exercise with blacks and whites. Again, let us say that you pull all the tests with a full-scale IQ score of exactly 105. Again, you examine the scores on the subtests. But this time the pattern of subtest scores is *not* the same for blacks and whites, even though the subtests add up to the identical overall score.⁸⁷ Despite identical overall scores, whites are characteristically stronger than blacks on the subtests involving spatial-perceptual ability, and blacks are characteristically stronger than whites in subtests such as arithmetic and immediate memory, both of which involve retention and retrieval of information.⁸⁸ As Jensen and Reynolds note, the pattern of subtest differences between whites and blacks differs sharply from the "no differences" result associated with SES. This directly contradicts the hypothesis that the B/W difference reflects primarily SES differences.⁸⁹ What accounts for the different subtest profiles? Jensen and Reynolds proceeded to demonstrate that the results are consistent with Spearman's hypothesis. Whites and blacks differ more on the subtests most highly correlated with *g*, less on those least correlated with *g*.

Since that initial study using the WISC-R, Jensen has been assembling studies that permit further tests of Spearman's hypothesis. He concluded from over a dozen large and representative samples of blacks and whites⁹⁰ that "Spearman's hypothesis has been borne out significantly by every study (i.e., 13 out of 13) and no appropriate data set has yet been found that contradicts Spearman's hypothesis."⁹¹ There appears to be no dispute with his summary of the facts. 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 positively correlated with the test's loading on *g*.⁹² The phenomenon seems peculiarly concentrated in comparisons of ethnic groups.

Jensen's most recent work on Spearman's hypothesis uses reaction time tests instead of traditional mental tests, bypassing many of the usual objections to intelligence test questions. Once again, the more *g*-loaded the activity is, the larger the B/W difference is, on average.⁹³ Critics can argue that the entire enterprise is meaningless because *g* is meaningless, but the hypothesis of a correlation between the magnitude of the *g*-loading of a test and the magnitude of the black-white difference on that test has been confirmed.^[94]

How does the confirmation of Spearman's hypothesis bear on the genetic explanation of ethnic differences? In plain though somewhat imprecise language: The broadest conception of intelligence is embodied in *g*. Anything other than *g* is either a narrower cognitive capacity or measurement error. Spearman's hypothesis says in effect that as mental measurement focuses most specifically and reliably on *g*, the observed black-white mean difference in cognitive ability gets larger.^[95] At the same time, *g* or other broad measures of intelligence typically have relatively high levels of heritability.^[96] 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.

Arguments Against a Genetic Explanation

The ubiquitous Arthur Jensen has also published the clearest evidence that the disadvantaged environment of some blacks has depressed their test scores. He found that in black families in rural Georgia, the elder sibling typically has a lower IQ than the younger.⁹⁷ The larger the age difference is between the siblings, the larger is the difference in IQ. The implication is that something in the rural Georgia environment was depressing the scores of black children as they grew older.^[98] In neither the white families of Georgia, nor white or black families in Berkeley, California, are there comparable signs of a depressive effect of the environment.

But demonstrating that environment can depress cognitive develop-

ment does not prove that the entire B/W difference is environmental, and in this lies an asymmetry between the contending parties in the debate. Those who argue that genes might be implicated in group differences do not try to argue that genes explain everything. Those who argue against them—Leon Kamin and Richard Lewontin are the most prominent—typically deny that genes have *anything* to do with group differences, a much more ambitious proposition.

CONFRONTING SPEARMAN'S HYPOTHESIS. If one is to make this case against a genetic factor on psychometric grounds, the data supporting Spearman's hypothesis must be confronted. There are two ways to do so: dispute the fact itself or grant the fact but argue that it does not mean what Jensen says it does.

The most searching debate about Spearman's hypothesis was conducted in a journal that publishes both original scholarly works and commentaries on them, *Behavioral and Brain Sciences*, where, in two separate issues in the latter 1980s, thirty-six experts in the relevant fields commented on Jensen's evidence.⁹⁹ A number of comments were favorable and provided further support for Jensen's conclusion. Others were critical, for reasons that varied from the philosophical (research into such hurtful issues is not useful) to the highly technical (were Jensen's results the result of varying reliabilities among the tests?). We summarize them in the notes, but the striking feature was that no commentator was able to dispute the empirical claim that the racial gap in cognitive performance scores tends to be larger on tests or activities that draw most on *g*.^[100]

Several years after the exchange on Spearman's hypothesis in *Behavioral and Brain Sciences*, Jan-Eric Gustafsson presented some data finding a considerably smaller correlation than Jensen and others do between *g* loading and B/W differences on a group of subtests.¹⁰¹ It is not clear why Gustafsson obtained these atypical results, but, as of this writing, they are still atypical. We have found no others for representative groups of blacks and whites. Our own appraisal of the situation is that Jensen's main contentions regarding Spearman's hypothesis are intact and constitute a major challenge to purely environmental explanations of the B/W difference.

CULTURAL EXPLANATIONS. Another approach has been taken by Jane Mercer, a sociologist and the developer of the System of Multicultural Pluralistic Assessment (SOMPA). Tests are artifacts of a culture, she argues, 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 IQ tests. Fewer books in the home means 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 Mercer accepts the generalizations that we discussed earlier about internal evidence of bias.¹⁰² She is 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, she argues, the effects of culture are more diffuse. Her argument may be seen as a variant of the "uniform background radiation" hypothesis that we discussed earlier.

Furthermore, she points out, strong correlations between home or community life and IQ scores are readily found. In a study of 180 Latino and 180 non-Latino white elementary school children in Riverside, California, Mercer 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 correlation between ethnic group and IQ among the children fell to near zero.¹⁰³

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? Recall that we pointed out earlier that controlling for socioeconomic status typically reduces the B/W difference by about a third. To the extent that parental socioeconomic status is produced by parental IQ, controlling for socioeconomic status controls for parental IQ. One obvious criticism of SOMPA is that it broadens the scope of the control variables to such an extent that the procedure becomes meaningless. After the correlations between the eight sociocultural variables and IQ are, in effect, set to zero, little difference in IQ remains among her ethnic samples. But what does this mean? The obvious possibility is that Mercer has demonstrated only that parents matched on IQ will produce children with similar IQs—not a startling finding.

Mercer points out that the samples differ on the sociocultural variables even after controlling for IQ. The substantial remaining correlations indicate that "important amounts of the variance in sociocultural characteristics [are] unexplained by IQ,"¹⁰⁴ evidence, she says, that they may be treated as substantially independent of IQ.^[105] But they are, in fact, not independent of IQ. They remain correlated. Her basic conclusion that "there is no justification for ignoring sociocultural factors when interpreting between-group differences in IQ" seems to us unchallengeable.¹⁰⁶ In the next chapter, we will present other examples of ethnic differences in social behavior that persist after controlling for IQ. But to conclude that genetic differences are ruled out by her analysis is unwarranted, because she cannot demonstrate that a family's sociocultural characteristics are independent of their IQ.¹⁰⁷

Scholars of Jensen's school point to a number of other difficulties with Mercer's interpretation. When she concludes that cultural diffusion explains the black-white difference, the data she uses show the familiar pattern of Spearman's hypothesis: The more a test loads on *g*, the greater is the B/W difference.¹⁰⁸ Why should cultural diffusion manifest itself in such a patterned way? Her appeal to sociocultural factors does not explain why blacks score lower on backward digit span than forward; why in chronometric tests, black movement time is faster, but reaction time slower, than among whites; or why the B/W difference persists on non-verbal tests such as the Ravens Standard Progressive Matrices. It is also not explained why, if the role of European white cultural diffusion (or the lack of it) is so important in depressing black test performance, it has been so unimportant for Asians.

A number of authors besides Mercer have advanced theories of cultural difference, often treated as part of the "cultural bias" argument but asserting in more sweeping fashion that cultures differ in ways that will be reflected in test scores. 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 non-material 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.”¹⁰⁹

The notes reference a variety of other authors who have made similar arguments.¹¹⁰ All, in different ways, purport to explain how large B/W 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.¹¹¹ 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.

THE FLYNN EFFECT. Indirect support for the proposition that the observed B/W difference could be the result of environmental factors is provided by the worldwide phenomenon of rising test scores.¹¹² 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 IQ scores often rose with every successive year after a test was first standardized. For example, when the Stanford-Binet IQ was restandardized in the mid-1930s, it was observed that individuals earned lower IQs 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.¹¹³ This meant that the average person could answer more items on the old test than the new test. Most of the change has been concentrated in the nonverbal portions of the tests.

The tendency for IQ scores to drift upward as a function of years since standardization has now been substantiated, primarily by Flynn, in many countries and on many IQ tests besides the Stanford-Binet.¹¹⁴ In some countries, the upward drift since World War II has been as much as a point a year for some spans of years. The national averages have in fact changed by amounts that are comparable to the fifteen or so IQ points separating whites and blacks in America. To put it another way, on the average, whites today may differ in IQ 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 15 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.

There is a further question to answer: Does a 15-point IQ difference between grandparents and their grandchildren mean that the grandchildren are 15 points *smarter*? Some experts do not believe that the rise is wholly, perhaps not even partly, a rise in intelligence but in the narrower skills involved in intelligence test taking per se;¹¹⁵ others believe that at least some of rise is in genuine intelligence, perhaps owing to the improvements in public education (by the schools and the media), health care, and nutrition. There is evidence that the rise in scores may be due to a contraction in the distribution of test scores in the population at large, with most of the shrinkage in the bottom half of the distribution.¹¹⁶ In large-scale studies of the Danish population, virtually all of the upward drift in intelligence test scores is accounted for by the rising performances of the lower half of the distribution.¹¹⁷ The data we presented earlier on the rise in SAT scores by American blacks are consistent with this story. In general, egalitarian modern societies draw the lower tail of the distribution closer to the mean and thereby raise the average.^[118] These findings accord with everyday experience as well. Whether one looks at the worlds of science, literature, politics, or the arts, one does not get the impression that the top of the IQ distribution is filled with more subtle, insightful, or powerful intellects than it was in our grandparents' day.

WI
mean
rise in
the fi
age /
now.
some
the c
econ
Fl
do w
telli
stab
the
do r
whi
acri

RA
wh
anc
dle
psy
he
ro
Th
th
th
sa
ho
or
b:
1:
e
g
r
t
e
I

Whatever we discover about the reasons for the upward drift in the mean of the distribution of test scores, two points are clear. First, a rapid rise in intelligence does not plausibly stretch far into either the past or the future. No one is suggesting, for example, that the IQ of the average American in 1776 was 30 or that it will be 150 a century from now.^[119] The rising trend in test scores may already be leveling off in some countries.¹²⁰ Second, at any point in time, it is one's position in the distribution that has the most significant implications for social and economic life as we know it and also for the position of one's children.^[121]

Flynn suggests that the intergenerational change in IQ has more to do with a shifting link between IQ scores and the underlying trait of intelligence than with a change in intelligence per se.^[122] Even so, the instability of test scores across generations should caution against taking the current ethnic differences as etched in stone. There are things we do not yet understand about the relation between IQ and intelligence, which may be relevant for comparisons not just across times but also across cultures and races.

RACIAL ANCESTRY. Just over 100 families with adopted children of white, black, and mixed racial ancestry are being studied in an ongoing analysis of the effects of being raised by white adopting parents of middle or higher social status.¹²³ This famous transracial adoption study by psychologists Sandra Scarr and Richard Weinberg is the most comprehensive attempt yet to separate the effects of genes and of family environment on the cognitive development of American blacks and whites. The first reports (when the children were about 7 years old) indicated that the black and interracial children had IQs of about 106, well above the national black average or the black average in Minnesota, where the samples were drawn. This result pointed to a considerable impact of the home setting on intelligence. However, a racial and adoptive ordering on IQ existed even in the first follow-up: The mean IQs were 117 for the biological children of white parents, 112 for the white adoptive children, 109 for the adopted children with one black and one white or Asian parent, and 97 for the adopted children with two black parents.^[124] Altogether, the data were important and interesting but not decisive regarding the source of the B/W difference. They could most easily have been squared with a theory that the B/W difference has both genetic and environmental elements in it, but, with considerable straining, could perhaps have been stretched to argue for no genetic influence at all.

A follow-up a decade later, with the children in adolescence, does not favor the no-genetics case.¹²⁵ The new ordering of IQ means was 109 for the biological children of white parents, 106 for the white adoptive children, 99 for the adopted children with one black parent, and 89 for the adopted children with two black parents.^[126] The mean of 89 for adopted children with two black parents was slightly above the national black mean but not above the black mean for the North Central United States. The bottom line is that the gap between the adopted children with two black parents and the adopted children with two white parents was seventeen points, in line with the B/W difference customarily observed. Whatever the environmental impact may have been, it cannot have been large.

Scarr and Weinberg continue to argue that the results are consistent with some form of mixed gene and environmental source of the B/W difference, which seems to us the most plausible conclusion.¹²⁷ But whatever the final consensus about the data may be, the debate over the Minnesota transracial adoption study has shifted from an argument about whether the environment explains all or just some of the B/W difference to an argument about whether it explains more than a trivial part of the difference.

Several smaller studies bearing on racial ancestry and IQ were well summarized almost two decades ago by Loehlin, Lindzey, and Spuhler.¹²⁹ They found the balance of evidence tipped toward some sort of mixed gene-environment explanation of the B/W difference without saying how much of the difference is genetic and how much environmental.^[130] This also echoes the results of Snyderman and Rothman's survey of contemporary specialists.

The German Story

One of the intriguing studies arguing against a large genetic component to IQ differences came about thanks to the Allied occupation of Germany following World War II, when about 4,000 illegitimate children of mixed racial origin were born to German women. A German researcher tracked down 264 children of black servicemen and constructed a comparison group of 83 illegitimate offspring of white occupation troops. The results showed no overall difference in average IQ.¹²⁸ The actual IQs of the fathers were unknown, and therefore a variety of selection factors cannot be ruled out. The study is inconclusive but certainly consistent with the suggestion that the B/W difference is largely environmental.

But dissenting voices can be heard in the academic world. For example, a well-known book, *Not in Our Genes*, by geneticist Richard Lewontin and psychologists Steven Rose and Leon Kamin, criticizes anyone who even suggests that there may be a genetic component to the B/W difference or who reads the data as we do, as tipping toward a mixture of genetic and environmental influences.¹³¹ How can they do this? Mostly by emphasizing those aspects of the data that suggest environmental influences, such as the correlations between the adopting parents' IQs or educational levels and the IQs of their black adopted children in the Minnesota study from the first follow-up (the book was published before the second follow-up). But they have nothing to say about the aspects that are consistent with genetic influence, such as the even larger correlations between the educational level of either the biological mothers or fathers and the IQs of their adopted-away black children.¹³² Although Lewontin, Rose, and Kamin do not say it in so many words, their argument makes sense if it is directed at the claim that the B/W difference is *entirely* genetic. It does little to elucidate the ongoing scientific inquiry into whether the difference has a genetic component.

We have touched on only the highlights of the arguments on both sides of the genetic issue. One main topic we have left untouched involves the malleability of intelligence, with two extremes of thought: that intelligence is remarkably unmalleable, which undercuts environmental arguments in general and cultural ones in particular, and that intelligence is highly malleable, supporting those same arguments. Because the malleability of intelligence is so critical a policy issue, it deserves a chapter of its own (Chapter 17).

RETHINKING ETHNIC DIFFERENCES

If the reader is now convinced that either the genetic or environmental explanation has won out to the exclusion of the other, we have not done a sufficiently good job of presenting one side or the other. It seems highly likely to us that both genes and the environment have something to do with racial differences. What might the mix be? We are resolutely agnostic on that issue; as far as we can determine, the evidence does not yet justify an estimate.

We are not so naive to think that making such statements will do much good. People find it next to impossible to treat ethnic differences

with detachment. That there are understandable reasons for this only increases the need for thinking clearly and with precision about what is and is not important. In particular, we have found that the genetic aspect of ethnic differences has assumed an overwhelming importance. One symptom of this is that while this book was in preparation and regardless of how we described it to anyone who asked, it was assumed that the book's real subject had to be not only ethnic differences in cognitive ability but the genetic source of those differences. It is as if people assumed that we are faced with two alternatives: either (1) the cognitive difference between blacks and whites is genetic, which entails unspoken but dreadful consequences, or (2) the cognitive difference between blacks and whites is environmental, fuzzily equated with some sort of cultural bias in IQ tests, and the difference is therefore temporary and unimportant.

But those are not the only alternatives. They are not even alternatives at all. The major ethnic differences in the United States are not the result of biased tests in the ordinary sense of the term. They may well include some (as yet unknown) genetic component, but nothing suggests that they are entirely genetic. And, most important, it matters little whether the genes are involved at all.

We have already explained why the bias argument does not readily explain the ethnic differences and also why we say that genes may be part of the story. To show why we believe that it makes next to no difference whether genes are part of the reason for the observed differences, a thought experiment may help. Imagine that tomorrow it is discovered that the B/W difference in measured intelligence is entirely genetic in origin. The worst case has come to pass. What difference would this news make in the way that you approach the question of ethnic differences in intelligence? Not someone else but *you*. What has changed for the worse in knowing that the difference is genetic? Here are some hypothetical possibilities.

If it were known that the B/W difference is genetic, would I treat individual blacks differently from the way I would treat them if the differences were environmental? Probably, human nature being what it is, some people would interpret the news as a license for treating all whites as intellectually superior to all blacks. But we hope that putting this possibility down in words makes it obvious how illogical—besides utterly unfounded—such reactions would be. Many blacks would continue to be smarter than many whites. Ethnic differences would continue to be dif-

ferences in means and distributions; they would continue to be useless, for all practical purposes, when assessing individuals. If you were an employer looking for intellectual talent, an IQ of 120 is an IQ of 120, whether the face is black or white, let alone whether the mean difference in ethnic groups were genetic or environmental. If you were a teacher looking at a classroom of black and white faces, you would have exactly the same information you have now about the probabilities that they would do well or poorly.

If you were a government official in charge of educational expenditures and programs, you would continue to try to improve the education of inner-city blacks, partly out of a belief that everyone should be educated to the limits of his ability, partly out of fairness to the individuals of every degree of ability within that population—but also, let it be emphasized, out of a hardheaded calculation that the net social and economic return of a dollar spent on the elementary and secondary education of a student does not depend on the heritability of a group difference in IQ. More generally: *We cannot think of a legitimate argument why any encounter between individual whites and blacks need be affected by the knowledge that an aggregate ethnic difference in measured intelligence is genetic instead of environmental.*

It is true that employers might under some circumstances find it economically advantageous to use ethnicity as a crude but inexpensive screen to cut down hiring costs (assuming it were not illegal to do so). But this incentive exists already, by virtue of the existence of a difference in observed intelligence regardless of whether the difference is genetic. The existence of the difference has many intersections with policy issues. The *source* of the difference has none that we can think of, at least in the short term. Whether it does or not in the long term, we discuss below.

If the differences are genetic, aren't they harder to change than if they are environmental? Another common reaction, this one relies on false assumptions about intelligence. The underlying error is to assume that an environmentally caused deficit is somehow less hard-wired, that it has less impact on "real" capabilities, than does a genetically caused deficit. We have made this point before, but it bears repeating. Some kinds of environmentally induced conditions can be changed (lack of familiarity with television shows for a person without a television set will probably be reduced by purchasing him a television set), but there is no reason to think that intelligence is one of them. To preview a conclu-

sion we will document at length in Chapter 17, an individual's realized intelligence, no matter whether realized through genes or the environment, is not very malleable.

Changing cognitive ability through environmental interventions has proved to be extraordinarily difficult. At best, the examples of special programs that have permanently raised cognitive ability are rare. Perhaps as time goes on we will learn so much about the environment, or so much about how intelligence develops, that effective interventions can be designed. But this is only a hope. Until such advances in social interventions come about, which is unlikely to happen any time soon, it is essential to grasp the point made earlier in the book: A short person who could have been taller had he eaten better as a child is nonetheless really short. The corn planted in the Mojave Desert that could have flourished if it had been planted in Iowa, wasn't planted in Iowa, and there's no way to rescue it when it reaches maturity. Saying that a difference is caused by the environment says nothing about how real it is.

Aren't genetic differences passed down through the generations, while environmental differences are not? Yes and no. Environmentally caused characteristics are by definition not heritable in the narrow technical sense that they do not involve genetic transmission. But nongenetic characteristics can nonetheless run in families. For practical purposes, environments are heritable too. The child who grows up in a punishing environment and thereby is intellectually stunted takes that deficit to the parenting of his children. The learning environment he encountered and the learning environment he provides for his children tend to be similar. The correlation between parents and children is just that: a statistical tendency for these things to be passed down, despite society's attempts to change them, without any necessary genetic component. In trying to break these intergenerational links, even adoption at birth has its limits. Poor prenatal nutrition can stunt cognitive potential in ways that cannot be remedied after birth. Prenatal drug and alcohol abuse can stunt cognitive potential. These traits also run in families and communities and persist for generations, for reasons that have proved difficult to affect.

In sum: If tomorrow you knew beyond a shadow of a doubt that all the cognitive differences between races were 100 percent genetic in origin, nothing of any significance should change. The knowledge would give you no reason to treat individuals differently than if ethnic differences were 100 percent environmental. By the same token, knowing

that the differences are 100 percent environmental in origin would not suggest a single program or policy that is not already being tried. It would justify no optimism about the time it will take to narrow the existing gaps. It would not even justify confidence that genetically based differences will not be upon us within a few generations. The impulse to think that environmental sources of difference are less threatening than genetic ones is natural but illusory.

HOW ETHNIC DIFFERENCES FIT INTO THE STORY

In any case, you are *not* going to learn tomorrow that all the cognitive differences between races are 100 percent genetic in origin, because the scientific state of knowledge, unfinished as it is, already gives ample evidence that environment is part of the story. But the evidence eventually may become unequivocal that genes are also part of the story. We are worried that the elite wisdom on this issue, for years almost hysterically in denial about that possibility, will snap too far in the other direction. It is possible to face all the facts on ethnic and race differences in intelligence and not run screaming from the room: That is the essential message.

This chapter is also central to the larger themes of the book, which is why we ask readers who have started with Part III to turn back to the Introduction and begin the long trek. In Part I, we described the formation of a cognitive elite. Given the cognitive differences among ethnic and racial groups, the cognitive elite cannot represent all groups equally, a statement with implications that we will develop in Part IV. In Part II, we described how intelligence is important for understanding the social problems of our time. We limited the discussion to whites to make it easier to think about the evidence without constantly having to worry about racism, cultural bias in the tests, or other extraneous issues.

The material in this chapter lets us proceed. As far as anyone has been able to determine, IQ scores on a properly administered test mean about the same thing for all ethnic groups. A substantial difference in cognitive ability distributions separates whites from blacks, and a smaller one separates East Asians from whites. These differences play out in public and private life. In the rest of Part III, we may now examine the relationship between social problems and IQ on a national scale.