

I.Q.

by **Richard Herrnstein**

- 1. If differences in mental abilities are inherited, and**
- 2. if success requires those abilities, and**
- 3. if earnings and prestige depend on success,**
- 4. then social standing will be based to some extent on inherited differences among people.**

TRUE? **FALSE?**

I.Q. tests and their like have become controversial, in spite of the hundreds of millions of them still given annually around the world. Especially in racially mixed urban centers of the United States—in New York, Los Angeles, Boston, Philadelphia—just their use in public schools evokes increasingly adamant protest. Because there are statistically reliable differences in I.Q. between whites and blacks, between the privileged and the underprivileged, and among various ethnic minorities, some influential people argue that the tests retard the liberalization of American society. Even the U.S. Supreme Court shares this view, as it recently proved by enjoining a firm from giving intelligence tests to potential employees. The Court ruled that since the menial jobs in question required virtually no intellectual distinction, the tests were serving to abet illegal discrimination, especially against blacks. Whatever one thinks of the merits of that decision or of the broader trend against mental testing, one may wonder why the issue itself has so suddenly arisen. Three landmark social documents of the past five years—by Daniel Patrick Moynihan, James S. Coleman, and Arthur R. Jensen—mark the critical stages as the concept of I.Q. has moved into its present embattled position.

Speaking on June 4, 1965, at Howard University, Lyndon B. Johnson warned that mere legislation, no matter how bold, was incapable of evening the score for American blacks. He alluded to the growing number of broken black

families as a sign of troubles to come, even as new laws and ground-breaking court decisions were supposed to reduce racial disparities. On this occasion, he promised renewed efforts to redress the inequities, leading off with a White House conference on civil rights that coming fall. The President did not acknowledge at the time that his facts and recommendations were garnered mainly from a confidential report by Moynihan, who had been Assistant Secretary of Labor from 1963 to early in 1965. The Moynihan report came to light only later that summer, after riots in Los Angeles had vividly confirmed the increasing discontent of blacks in the cities. According to Moynihan, American blacks, suffering under the bitter, supposedly emasculating legacy of slavery, had evolved a matriarchal family structure which was seriously out of line with the rest of American society. Black children, and the adults they matured into, were consequently at a disadvantage in our primarily patriarchal culture. Much to the surprise of the government, the Moynihan report was vehemently rejected at the civil rights conference that fall. The trouble, said civil rights spokesmen, was not in the black family but in white racism. Give \$100 billion to clear up the city slums, said A. Philip Randolph, and the black problem will take care of itself. The Moynihan report, which Johnson and his advisers had seen as a significant step forward in racial understanding, was scorned as a new, “subtle” form of racism.

The Johnson Administration showed a surer grasp of racial dialectics in handling the Coleman report about a year later. This scholarly empirical study, conducted under the guidance of a Johns Hopkins professor of sociology, James S. Coleman, was originally authorized by the Civil Rights Act of 1964. The U.S. Office of Education was given two years to assess the inequities in American public schools, and two years, 4000 schools, 60,000 teachers, and 605,000 students later, Commissioner of Education Harold Howe announced the main findings. But the announcement was made the Friday afternoon in July preceding the holiday weekend, probably, as Senator Abraham Ribicoff (Democrat of Connecticut) later noted angrily, in an effort to minimize public attention. No doubt the government was uneasy about the findings. Blacks lagged behind whites in scholastic achievement at every grade level from first to twelfth, and the differences increased with age. Ordinarily, one might blame the general inferiority of segregated black schools for that difference, but the Coleman study sought without success any clear effect of school quality on scholastic achievement for white children. If schools themselves deserve the blame for the poorer performance of blacks, then why shouldn't the whites be similarly affected? The answer seemed to be that there was some other difference between the white and black children besides their schools, and the tentative, guarded, and little-publicized hypothesis of Commissioner Howe and Secretary of Health, Education and Welfare John W. Gardner was that the difference was in the cultural surroundings at home—a touchy subject, as the reaction to the Moynihan report had amply shown.

Both the Moynihan and Coleman reports grappled with the idea that something within the black community itself was holding back its economic and educational advance. Neither report denied the clear evidence that racist customs and even laws were in large part responsible for the lag. But both reports noted that, for reasons not wholly understood, the removal of external barriers such as racist customs and laws did not always bring the promised improvement in economic and educational condition, presumably because of internal barriers—for example, family structure or cultural ambience. Such a presumption made both reports intensely unwelcome to civil rights interests. Understandably, anything that transferred the burden of melioration from whites to blacks was immediately suspect as racist. Both Moynihan and Secretary Gardner, the latter commenting on the Coleman report, blamed the cultural history of American blacks—which is to say, they blamed slavery—for those internal barriers. But the third and most controversial document—Berkeley professor Arthur R. Jensen's article published in the Harvard Educational Review in the winter of 1969—faced head-on the possibility that blacks and whites differ in inherited intelligence. This difference, which shows up as the average difference in their I.Q.'s, may be the extra factor which gives whites a statistical advantage in economic and educational competition in certain settings. Although

Jensen did not assert that this had been proved, his consideration of it provoked so violent a reaction that the earlier reactions to Moynihan and Coleman seem polite by comparison. SDS was on the streets of Berkeley almost immediately with bullhorns blaring, "Fight racism! Fire Jensen!" Jensen's classes had to meet clandestinely to avoid repeated disruptions by outraged activists. Some of Jensen's colleagues at the University of California tried, unsuccessfully, to have him censured. His hate mail was voluminous. In apparent panic over the vehemence of the outcry, the Harvard Educational Review refused to sell reprints of the article to anyone (including Jensen) until they could be bound together with a number of criticisms of Jensen's arguments.

The Jensen report (as this article has come to be mis-called) dealt with intelligence and inheritance in general, not only with racial questions. Other writers in the past four centuries—from Thomas Hobbes to Konrad Lorenz—have agonized over the complex and fascinating interplay of nature and nurture in shaping man's psyche. It is only lately in America that public discussion requires physical, not to mention intellectual, courage, for the subject is close to taboo. But The Atlantic believes that it is not only possible but necessary to have public discussion of important, albeit painful, social issues. The subject of intelligence is such an issue—important because social legislation must come to terms with actual human potentialities, painful because the actualities are sometimes not what we vainly hope.

—The Editors

The measurement of intelligence forced its way into America's public consciousness during World War I, when almost two million soldiers were tested by the Army and categorized as "alpha" and "beta," for literates and illiterates respectively. The lasting effect of that innovation has not been the surprise at learning that the average American soldier had an intelligence equal to that of a thirteen year old, or that artillery officers were substantially brighter than medical officers, or any of the myriad other statistical curiosities. Even if those facts are still as true as they were in 1918, the lasting effect has been the mere use of the tests and their serious consideration by responsible people. For intelligence tests, and the related aptitude tests, have more and more become society's instrument for the selection of human resources. Not only for the military, but for schools from secondary to professional, for industry, and for civil service, objective tests have cut away the traditional grounds for selection—family, social class, and, most important, money. The traditional grounds are, of course, not entirely gone, and some social critics wonder if they do not lurk surreptitiously behind the scenes in our definition of mental ability.

But at least on the face of it there is a powerful trend

The I.Q. test, it is said with fervor, is used by the establishment to promote its own goals and to hold down the downtrodden.

toward “meritocracy”—the advancement of people on the basis of ability, either potential or fulfilled, measured objectively.

Lately though, the trend has been deplored, often by the very people most likely to reap the benefits of measured intellectual superiority. More than a few college professors and admissions boards and even professional testers have publicly condemned mental testing as the basis for selection of people for schools or jobs. The I.Q. test, it is said with fervor, is used by the establishment to promote its own goals and to hold down the downtrodden—those non-establishment races and cultures whose interests and talents are not fairly credited by intelligence tests. These dissenting professors and testers are naturally joined by spokesmen for the disadvantaged groups. We should, these voices say, broaden the range of humanity in our colleges (to pick the most frequent target) by admitting students whose low college entrance examination scores might otherwise have barred the way. For if the examinations merely fortify an arbitrarily privileged elite in its conflict with outsiders, we must relinquish them. The ideals of equality and fraternity must, according to this view, take precedence over the self-interest of the American-Western European middle class.

The issue is intensely emotional. It is almost impossible for people to disagree about the pros and cons of intelligence testing and long avoid the swapping of oaths and anathema. Yet should not the pros and cons be drawn from facts and reason rather than labels and insults? For example, is it true that intelligence tests embody only the crass interests of Middle America, or do they draw on deeper human qualities? Is the I.Q. a measure of inborn ability, or is it the outcome of experience and learning? Can we tell if there are ethnic and racial differences in intelligence, and if so, whether they depend upon nature or nurture? Is there only one kind of intelligence, or are there many, and if more than one, what are the relations among them? If the tests are inadequate—let us say, because they overlook certain abilities or because they embody arbitrary cultural values—how can they be improved? For those who have lately gotten their information about testing from the popular press, it may come as a surprise that these hard questions are neither unanswerable nor, in some cases, unanswered. The measurement of intelligence is psychology’s most telling accomplishment to date. Without intending to belittle other psychological ventures, it may be fairly said that nowhere else—not in psychotherapy, educational reform, or consumer research—has there arisen so potent an instrument as the objective measure of intelligence. No doubt intelligence testing is imperfect, and may even be in some sense imperfectible, but there has already been too much success for it to be repudiated on technical grounds alone. If intelligence testing is to change, it must change in light of what is known, and more is known than most might think.

Mental testing was one of many responses within psychology to Darwin’s theory of evolution. In fact, the connection here is intimate and direct, for the idea of measuring mental ability objectively was first set forth by Francis Galton, the younger cousin of Charles Darwin. Far more versatile (perhaps smarter) than his great cousin, Galton was a geographer, explorer, journalist, mathematician, eugenicist (he coined the term), and articulate essayist. In 1869, just a decade after Darwin launched modern biology with the *Origin of Species*, Galton published *Hereditary Genius*, which applied evolutionary thinking to the question of intellect. Galton noted, first, that men varied greatly in their intellectual capacity and, second, that various kinds of excellence run in families, suggesting that the basis of intelligence may be inherited. Going back through British history, Galton found that judges, statesmen, prime ministers, scientists, poets, even outstanding wrestlers and oarsmen tended, for each kind of endeavor, to be related by blood. The eminent families of Great Britain were taken as evidence of superior human strains, comparable to the natural biological variations that figure so prominently in the doctrine of evolution. Today, our sensitivity to the role of the environment (not to mention such mundane complications as money and family connections) make us skeptical of his evidence. Nevertheless, in the first flush of Darwinian social theorizing, Galton called for constructive change. The inheritance of human capacity implied “the practicability of supplanting inefficient human stock by better strains,” and led him “to consider whether it might not be our duty to do so by such efforts as may be reasonable, thus exerting ourselves to further the ends of evolution more rapidly and with less distress than if events were left to their own course.”

Galton was not much more content with the genealogical approach to mental ability than are we today. Within a few years, he was trying to test mental ability directly, but the problem was how to do it. In 1882, Galton set up a small laboratory in a London museum where people could, for a fee, have their hearing, vision, and other senses tested. Galton knew that mental defectives—idiots and imbeciles—often lacked sensory acuity, and he guessed that there might be a reasonably consistent relation between intelligence and sensory keenness in general. As it turned out, his hunch was wrong, or at least not right enough to be useful as a way of testing on a large scale.

Galton was soon just one of many scientists searching for a practical intelligence test, with no one much worried at this point about the ultimate definition of intelligence. Intuition and common sense set the standards as the few simple measures of sensory acuity gave way to a host of tests, some sensory and others drawing on other psychological processes. An American psychologist named James McK. Cattell coined the phrase “mental test” in 1890 in an article recounting his studies at the University

Whatever intelligence is, it varies from person to person.

of Pennsylvania on the mental abilities of students. In addition to simple sensory function, Cattell measured color discrimination, time perception, accuracy of hand movement, and memory; and he collected descriptions of imagery. People no doubt differed, but it was hard to know what to make of the differences. By the mid-1890s, testing had attracted so much attention that professional organizations began taking note of it. The newly founded American Psychological Association formed a committee in 1895 “to consider the feasibility of cooperation among the various psychological laboratories in the collection of mental and physical statistics”; in 1896 the American Association for the Advancement of Science instructed a committee of its own “to organize an ethnographic survey of the white race in the United States.” The quotations in both cases are Professor Cattell’s words; he was a member of both committees and was determined that the ethnographic survey for AAAS include some of APA’s mental (and physical) tests.

For all of the ferment, it was not yet certain that anything useful was brewing. There was spirit and energy in abundance, but there were as yet no indisputably good tests. It took the work of a French psychologist named Alfred Binet to make intelligence testing practical. In a key article written in 1895, Binet and his junior collaborator, Victor Henri, argued for mental testing based not on sensory or motor functions but on the psychological processes thought to be involved in intelligence. Instead of supposing that being smart is the outcome of having keen senses or speedy reactions, Binet argued that intelligence operates at its own level and that, therefore, a proper test must engage the person at that very level. As for what such tests might be, Binet, like everyone else in 1895, was just guessing. The article suggested a variety: tests of memory, mental imagery, imagination, attentiveness, mechanical and verbal comprehension, suggestibility, aesthetic appreciation, moral sensibility, the capacity to sustain muscular effort, and visual judgment of distance.

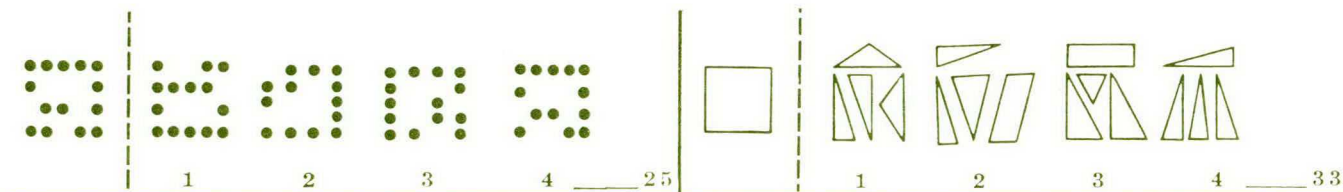
Binet criticized his contemporaries for their preoccupation with sensory and other simple processes, which, although fulfilling their desire for exactitude in measurement, had sacrificed the still more salient need for relevance. For Binet, exactitude was secondary. His pragmatism directed him to tests that sorted people out—for, whatever intelligence is, it varies from person to person. The sensory data did not distinguish among people as sharply as intuition required for a test of intelligence. Binet committed himself to seeking the tests that would do so, which was an undertaking that occupied the rest of his life. In the following ten years, Binet and his collaborators worked on mental testing at the psychological laboratory of the Sorbonne, using as their subjects mainly children from the schools of Paris and its suburbs.

The use of children was a happy accident, for it focused attention on the chronology of intelligence. Of all the

countless ways one may want to distinguish between smarter and duller people, it may not seem especially insightful to choose the simple fact that during the first fifteen or so years of life, age confers intelligence (on the average). Thus, if an intellectual task sorted children according to their age, then it might properly be included in an intelligence test. In one experiment, for example, Binet tested over five hundred schoolchildren by reading them a sentence and then asking them to write down as much of the sentence as they could remember. Between the ages of nine and twelve (the ages tested), each successive grade of student did better, albeit slightly, than the grade younger. From this, Binet knew that the “sentence-reproduction test” could be taken as one measure of mental capacity. And knowing that, he could say that if two children of equal age differed in their sentence-reproduction scores, they were to some degree different in intelligence. One such test was, however, far from a usable measure of general intelligence, as Binet well knew.

As the years passed, Binet and others stocked a rich store of norms and measures of mental ability, based on many tests of many children. Even Binet’s own two daughters were the subject of intensive study, culminating in a book called *The Experimental Study of Intelligence* (1902), in which the vital psychological facts about the teen-age girls were expressed as scores on their father’s tests of word-writing speed, mental imagery, sentence completion, and so on. It was to Binet, therefore, that the Minister of Public Instruction turned in the fall of 1904 when he wanted a better way to spot subnormal children in the Parisian schools. The children were to be put into special schools where they could be helped, but the first problem was to find them. If mental tests were any use at all, here was a task to prove it. Binet and his psychiatrist collaborator, Theodore Simon, decided to use a series of tests graded in difficulty, first standardized on normal children of various ages.

The idea of using equivalent age as the measure of intelligence was obvious only after Binet, not before him, for it was one of those rare and elegant turns that make for historic innovation. Here were some tests that distinguished between children of different ages, on the average. However, at each age some children did better than their exact chronological peers. Those children, he had found, were judged by teachers to be bright or gifted. Conversely, other children did worse than their peers and were judged to be dull. Hence, if all one knows about a child is that he outperforms his age peers, he can still be assumed to be bright. If his performance matches his age, he is probably an average child in intelligence. And if he underperforms, he is probably dull. As Binet well knew, the chronological approach to intelligence finessed the weighty problem of defining intelligence itself. He had measured it without having said what it was. It took a while to know whether the sleight of hand had in fact



yielded a real intelligence test or just an illusion of one.

For their first practical venture, Binet and Simon drew up a progression of thirty tests covering the range of mental capacity. At the very bottom, the examiner simply noted eye-head coordination as a lighted match was moved across the field of vision; thence he observed the making of grasping movements, imitating gestures, the following of instructions to touch various parts of the body, the naming of familiar objects, repeating sentences, arranging identical-looking objects in order of weight, constructing sentences to include three given words ("Paris," "gutter," "fortune"); and finally the ability to distinguish between abstract words such as "liking" and "respecting." After some preliminary trials, Binet and Simon gave their test to about fifty normal children between the ages of three and eleven, thereby establishing the cutoffs for each age. Finally, using children already diagnosed by standard clinical procedures to be idiots, imbeciles, and morons, they found the corresponding criteria for mental disability in their series of tests.

Is a retarded child really the equal of a normal child at a younger age? For example, the average five year old passed the first fourteen tests, while the upper limit for an imbecile was to pass the first fifteen tests whatever age he was. Anyone who passed more was not an imbecile. Was Binet saying that a twelve-year-old imbecile precisely equals a slightly brighter-than-average five year old? The answer is no, for Binet specifically denied the charge. The imbecile, he said, is "infirm," the five year old is healthy, and their mental processes are in some respects different, even if the difference is not captured by his test. Nevertheless, the test did its job, for a twelve year old who tested at the five-year-old level was, indeed, retarded, while a five year old who did so was not (or at least did not seem to be at that time). As always, Binet's approach was doggedly pragmatic and empirical. He was picking out the retardates with his test more quickly, cheaply, and for all anyone knew, more accurately than ever before. The social benefits were self-evident.

The Binet-Simon test was put into use immediately and was criticized as quickly for this or that item. But criticism was corrective, for in showing that some item was not, for example, distinguishing between three and four year olds, the critic was opening the test to improvement. An ineffective item could be dropped, a useful one added, without in the least altering the kernel idea, which was to measure intelligence by a graded series of tasks ("stunts," Binet often called them). The tests and the criticisms were rooted in actual experience with ever-growing numbers of children, adding greater and greater empirical stability to the results. In America, Great Britain, Belgium, Italy, Germany, and elsewhere, the tests were being used and perfected. In a cheering counterexample to Gresham's gloomy law, good test items tended to drive out bad ones, and the better the test in sorting out children, the more it

was used and improved. In 1908 Binet and Simon published a much-revised series of tests, to be used for rating children in general, not just retarded children. In 1911 the final Binet-Simon scale came out; it was Binet's last work, for he died that year at the age of fifty-four. But the evolution of testing continued unabated and still does.

In the 1911 version, there were five problems which the average child of each age could or could not solve. Here, for example, are the five items for the six-year level:

1. Distinguish between morning and afternoon.
2. Define familiar objects in terms of use.
3. Copy a diamond shape.
4. Count thirteen pennies.
5. Distinguish between ugly and pretty faces.

And here are the five problems for the average ten year old:

1. Arrange five blocks in order of weight.
2. Draw two designs from memory.
3. Criticize absurd statements.
4. Answer comprehension questions.
5. Use three words in not more than two sentences.

A child who passes all the tests up to and including those for six year olds and none beyond has a "mental age" of six, whatever his actual chronological age. Suppose, however, that he passes all the tests up to but not including the six-year level, and then passes only three at the six-year level and one at the seven-year level. His mental age is credited with .2 additional years for every item he passes beyond the level where he has passed them all. This child's mental age would be $5 + .6 + .2$, or 5.8 years of mental age. If his chronological age were six years, he would be slightly below average; if five years, somewhat above.

Binet did not come up with the "intelligence quotient" (I.Q.) itself; this fell to the German psychologist William Stern to do soon thereafter. Stern saw that a child who is one year behind at the age of six is more retarded than a child who is one year behind at the age of thirteen. It is the *relation* between mental and chronological age that matters, not just their *difference*, and this relation is best expressed by the ratio of the two numbers. To get the I.Q., divide mental age by chronological age and multiply by 100 to get rid of the decimals. Thus, a six-year-old child who comes through with a mental age of nine is in these terms as bright as an eight year old with a mental age of twelve, both having the impressive I.Q. of 150.

The I.Q. of 100 divides the population into two roughly equal groups. This is not a fact of nature but an outcome of how the tests were made. Binet and his successors picked and chose until they found items that the average child at each age could just pass, thus assuring that the average child's mental age equals his chronological age and his I.Q. 100. The idea of a mental age assumes that mental growth is accumulative and consecutive, so that a child who has mastered the items at a given age level one year will (barring disease or trauma) continue to do at

least that well as he ages. In this case nature, not the test-makers, meets the condition. At each age during childhood we can do intellectually what we have done before, adding competence rather than replacing it. Binet's idea for mental testing would not have worked for grubs and caterpillars, which appear to lose their grasp of burrowing and cocoon spinning as they become competent at flight. In other respects too, Binet was fruitfully combining nature and artful design in his tests. Items on the test were included only if some children were ahead of their age in solving them, some behind, but the largest number were neither. Overall, the spread of performance conformed to the bell-shaped curve that statisticians call "normal," with about as many superior children as inferior, but with most crowding around the average.

Binet's ideas took hold powerfully and quickly. It was not only in France that the average eight-year-old child could just barely repeat accurately five digits read to him, for the Binet scale was readily exported to Belgium, Great Britain, America, Italy, and so on. The remarkable exportability of the tests was probably the first convincing argument for their soundness. Items that drew on bits of specific, seemingly arbitrary knowledge crossed national and linguistic boundaries as easily as the fundamental tests of memory and reasoning. It could be relied upon, for example, that the average nine year old would be able to name in order the months of the year. What does this say about the I.Q.? Would we downgrade a Papuan child, raised in New Guinea, if he could not name the months? Clearly not, if his language had no such names or had some different scheme for cutting up the year. Some of the items on a test are specific to a culture, but that does not make them poor items. A given test is only for people drawn from the same general population that the test was standardized on. Even if it is hard to locate the precise boundaries of this general population, a useful intelligence test should incorporate at least some of the material of a culture, or it may miss gauging the child's ability to assimilate his surroundings. Virtually every child grows up in some culture or another, and his intelligence score (if that concept is to retain its ordinary meaning) must reflect his sensitivity to it. The Papuan child cannot sensibly be tested on a Western intelligence test. He would do poorly, but he would also do poorly in most other contacts with Western society. It would not mean that he was not intelligent. It would mean only that he was not meeting the underlying conditions of the test, which assume that he has been drawn from the standardizing population. Analogously, a child who gets a very high I.Q. after being drilled by parents or teachers on test items is probably not all that bright, and for the same reason. Like any other instrument of measurement, the I.Q. test must be used according to the directions. One does not use an oral thermometer after eating hot soup or sucking on ice cubes—not if one wants to know one's temperature. One may have a fever with a

cool mouth, but the thermometer will not reveal it. So, the Papuan child may be bright or dull or average, but only a test standardized in his cultural environment can show which. It is not that "intelligence" itself is peculiarly European or North American, even if the instrument for gauging it is.

A person's I.Q. is a different sort of fact about him than his height or his weight or his speed in the hundred-yard dash, and not because of the difference between physical and mental attributes. Unlike inches, pounds, or seconds, the I.Q. is entirely a measure of relative standing in a given group. No such relativism is tolerated for the conventional measures. Gulliver may have looked like a giant in Lilliput and a mite in Brobdingnag, but he was just about 70 inches tall wherever he went. Relativism is tolerated for the I.Q. because, first of all, we have nothing better. If the testers came up with something like a platinum yardstick for mental capacity, it would quickly displace the I.Q. But more than this can be said for the I.Q. Because the group with which a child is implicitly compared is effectively the entire population of Western society, there is great stability to the comparison. The I.Q. gives one's standing among the people with whom one will live. And if it can be assumed that so large a sample of mankind is reasonably representative of the whole, then a relative measure is quite informative. An I.Q. of 100 would then indicate average intelligence, compared to people in general and not some small group; an I.Q. of 150 would denote high intelligence, and so on.

At around adolescence, people seem to stop acquiring new intellectual powers, as distinguished from new information or interests. For example, immediate memory span grows until the age of fifteen, but not thereafter. The average person can repeat seven digits at fifteen or at fifty. Other items in the Binet scale similarly level off at about the same age. Thus, if one were to continue calculating I.Q. in the same way, dividing a fixed mental age by a growing chronological age, one's score would plummet, reaching (for the average person) I.Q. 50 at about the age of thirty and I.Q. 25 at the age of sixty (assuming that the mental age is stuck at fifteen). To avoid such nonsense, some other measure of relative standing is often used for adults. Thus, instead of saying that a man has an I.Q. of 130, say instead that he tests higher than 96 percent of his peers, and then define the peer group. It can be all American adults, or Caucasians, or college graduates, or members of the United Auto Workers or the League of Women Voters. In fact, since the I.Q. is itself standardized on groups of peers (usually children), it and the percentile score are directly and simply translated one into the other.

Binet invented the modern intelligence test without saying what intelligence is. At first he was trying to sort out the mental defectives; later he was trying to rate all the children—defective, average, or superior. Some rough-and-ready notion of intelligence lurked in the background—

At around adolescence, people seem to stop acquiring new intellectual powers.

having to do with mental alertness, comprehension, speed, and so on—but he was not forced to defend an abstract definition in order to sell the idea of his test to the world. Instead, he could point to how well the test worked. Rarely did a bright child, as judged by the adults around him, score poorly, and rarely did a poor scorer seem otherwise bright. Occasionally a child would do worse than expected on the test because a teacher had confused obedience with brightness, or better than expected when rebelliousness had been mistaken for stupidity, but in general most children ended up about where they were expected to. The value of the test was that it gave an objective assessment about a child in an hour or so, and any trained technician could administer it. With the test as a yardstick, children who knew no one in common could be directly compared, for whatever purpose.

But is intelligence really an attribute, like height, that can be expressed in a single number? Even granting that I.Q. is a measure only of relative standing, can relative standing be given in a single number? Is Jimmy really altogether brighter than Johnny if his I.Q. is higher? Perhaps Jimmy is brighter as regards A,B,C, and D, but Johnny has him beaten on E, F, and G. Even Binet admitted that intelligence was not just one thing; otherwise his labors in creating a test would have been far easier. Once, when he was speculating about the nature of intelligence, Binet mentioned the attributes of directedness, comprehension, inventiveness, and critical capacity, which he thought may vary somewhat independently from person to person. Usually, however, he was too busy with his practical goals to dwell on hypotheses.

Even as Binet was developing the first intelligence scale, others were grappling with the conceptually tougher problem of the structure of intelligence. The story of the key mathematical discoveries would be out of place here, but the highlights may be worth noting. An Englishman named Charles Spearman resigned a commission in the British Army after serving in the Boer War and set to work on the problem. Taking the intercorrelations between scores on simple mental tests as his basis, he concluded that there was a “universal” intellectual capacity—which he labelled “g” for “general”—plus a host of minor, unrelated capacities of no great scope. The universal factor, he said, permeated all intellectual activity, while the others were variously absent or present in any given task. To be smart, for Spearman, mainly meant having lots of g. Although he had some evidence for this theory, it did not endure even for Spearman, who revised it after a decade or so. Nevertheless, his mathematical procedures were an essential link between Francis Galton’s formulas for assessing correlation and the vastly more complex methods of “multiple factor analysis,” which is the contemporary term.

Following Spearman, the next big step was taken by L. L. Thurstone, an American electrical engineer who left a job in Edison’s laboratory in East Orange, New Jersey, to work on psychological measurement. A long and illustrious career, covering the measurement not only of intelligence but also of attitudes, personality, sensory capacity, motivation, and the learning process was the result. For intelligence, Thurstone subdivided Spearman’s general factor, g, into a set of Primary Mental Abilities (PMA): spatial visualization, perceptual ability, verbal comprehension, numerical ability, memory, word fluency, and reasoning (inductive and deductive). These are just verbal labels tagged on at the end of a mathematical procedure that really has no verbal labels in it. It would be more precise (if less informative) to say that Thurstone found evidence for seven or eight separate factors or aspects of intelligence, and to leave it at that. With more powerful mathematics and more abundant data, Thurstone’s successors have teased out new factors. Like nuclear physics with its proliferation of elementary particles, the study of intelligence has suffered from its riches. Now there are experts who find evidence of over one hundred components in intelligence, and there is no sign of a limit.

Thurstone noted some intercorrelations among the Primary Mental Abilities. People who excelled, for example, in verbal comprehension were often high in word fluency. Other constellations also kept turning up. Such correlations among the factors themselves could signify that mental abilities are hierarchical, arranged in layers. At the very top, there may be a general intellectual power, like Spearman’s g, pervading all mental activity. To be smart means having the power in abundance, to be stupid means having a shortage, so that all of Thurstone’s PMA’s will be to some degree correlated. At the next level down, the PMA’s break into clusters involving either verbal abilities or numerical or logical abilities. Then there are the separate PMA’s themselves, which vary somewhat independently despite their intercorrelations. In addition to being generally bright or stupid or average, people are verbal, numerical, imaginative, and so on. People can be so strong in one factor or another that they excel in some areas without any special abundance of g. And, inversely, some people may be so poorly endowed in one or the other factors that they appear occasionally incompetent, notwithstanding substantial g. Although the hierarchy seems like a plausible theory of intelligence, it will remain hypothetical until the experts agree on its specific features—which has yet to happen.

Even at best, however, data and analysis can take us only so far in saying what intelligence is. At some point, it becomes a matter of definition. For example, we would reject any intelligence test that discounted verbal ability or logical power, but how about athletic prowess or manual dexterity or the ability to carry a tune or qualities of heart and character? More data are not the final answer, for at

bottom, subjective judgment must decide what we want the measure of intelligence to measure. So it is for all scales of measurement—physical as well as psychological. The idea of measuring length, weight, or time comes first; the instrument comes thereafter. And the instrument must satisfy common expectations as well as be reliable and practical. In the case of intelligence, common expectations center around the common purposes of intelligence testing—predicting success in school, suitability for various occupations, intellectual achievement in life. By this standard, the conventional I.Q. test does fairly well. The more complex measures, such as Thurstone's PMA's, add predictive power that is sometimes essential. As for what intelligence "really" is, the concept still has ragged edges where convenience and sheer intuition set boundaries that will no doubt change from time to time. The undisputed territory has, however, become formidable.

Most of us get our first, sometimes our only I.Q. test in school; the predictive power of the I.Q. is encountered first in our school grades; our teachers know our I.Q.'s even when our parents (let alone we ourselves) do not. But for all these connections, I.Q. and education are only correlated, not identical. First of all, there is the fact of variability: at each level of education, the I.Q.'s span a broad range, and at each level of I.Q. among adults, the amount of education completed also spans a broad range. Moreover, school grades show the effect of the environment more than I.Q.'s. And, finally, the correlation between I.Q. and schooling shows up even when the I.Q. is obtained from six year olds just starting school. Of course, once a child is known to have a high or low I.Q., he may live up, or down, to his teachers' expectations, but even granting that complication, the I.Q. could hardly predict how much schooling there is going to be in someone's life if it were itself just a result of schooling.

The discrepancies between I.Q. and school grades are instructive, because they follow a definite pattern. It is not just that the I.Q. is not an exact predictor of grades, but that children with low I.Q.'s almost always do poorly in school, while children with high I.Q.'s cover the range from excellent down to poor. For schoolwork, as for many other correlates of the I.Q., intelligence is necessary but not sufficient. It is as if a high I.Q. offers merely the opportunity for scholastic achievement, but something more is needed to exploit it. We can guess what the something more might be—interest, emotional well-being, energy—but we do not know. Other activities that are correlated with I.Q.—such as success in business—also seem to call on something more, although perhaps not the same extras as good schoolwork does. No doubt it takes physical strength and stamina to be a champion athlete, but for many sports it takes some intelligence as well. To be a successful actor

may take a good appearance or voice, but no doubt also intellect. The examples could be multiplied almost endlessly. I.Q. seems to be the *sine qua non* for an extraordinary variety of successes, but for virtually nothing practical is it the sole requirement.

But still, what is it? Even if it is not just schooling, may it not be a cryptic index of membership in the middle and upper classes, as many critics argue? We often hear that both I.Q. and successful education, and all the other correlates, follow from the more basic fact of social origin. To this criticism there is no short and simple answer. The correlation between I.Q. and social class (usually defined in terms of occupation, income, and patterns of personal association) is undeniable, substantial, and worth noting. A cautious conclusion, based on a survey of the scientific literature, is that the upper class scores about thirty I.Q. points above the lower class. A typical member of the upper class gets a score that certifies him as intellectually "superior," while a typical member of the lower class is a shade below average (that is, below I.Q. 100). Precise values cannot be taken too literally, for they depend on somewhat arbitrary definitions of social class and on which particular I.Q. test is used, but the basic finding is beyond dispute. Depending on whether one is for or against testing, one will see this class difference as a weakness either in the intellect of the underprivileged or in the tester's definition of intelligence. But in either case, there is no basis for assuming that *no* poor people have high I.Q.'s. On the contrary, many members of the lower class must have superior I.Q.'s, notwithstanding the low overall average. Recall that, by design, there are as many people above I.Q. 100 as below. In contrast, the social scale is definitely lopsided, with many more at the bottom than at the top even in affluent America. Only about 10 percent of our people meet the criteria for the upper and upper-middle classes, while about 65 percent are in the working class and below, with the remainder in between. But only 50 percent of the people have subnormal (below 100) I.Q.'s. And so, there must be at least 15 percent of our population in the bottom classes with supranormal (above 100) I.Q.'s.

It is one thing to note the correlation between social class and I.Q. but something else to explain, or even interpret it. It does not prove that the I.Q. is caused by social class, any more than it proves the reverse—that social class is caused by I.Q. More information is needed to sort out the possibilities. Since a family's social standing depends partly on the breadwinner's livelihood, there might be a further correlation between I.Q. and occupation. A large sample of enlisted men in the Air Force in World War II, drawn from seventy-four different civilian occupations, revealed in detail the expected I.Q. differences. Here are some of the findings, culled from a study by T. W. Harrell and M. S. Harrell published in 1945 in a periodical called *Educational and Psychological Measurement*:

A cautious conclusion is that the upper class scores about thirty I.Q. points above the lower class.

Rank in list of 74 occupations	Civilian occupation	Average I.Q.
1	accountant	128.1
5	auditor	125.9
10	draftsman	122.0
15	sales manager	119.0
20	clerk-typist	116.8
25	radio repairman	115.3
30	laboratory assistant	113.4
35	musician	110.9
40	sales clerk	109.2
45	power lineman	107.1
50	riveter	104.1
55	bartender	102.2
60	molder	101.1
65	baker	97.2
70	lumberjack	94.7
74	teamster	87.7

Each occupation has a range of I.Q.'s: not-so-bright accountants and very bright bakers are far from unknown. But just as for good grades in school, a high I.Q. is necessary for some occupations, even if it is not sufficient. For example, among the seventy-four civilian occupations that turned up in the group, public relations proved to have the fourth highest average I.Q., with the top I.Q. an impressive 149. The top truck driver also registered 149, but truck drivers averaged sixty-seventh in the list of occupations, close to the bottom with lumberjacks and teamsters. The lowest PR man had an I.Q. of 100, while the dullest truck driver tested an almost unbelievable 16—essentially no tested intelligence at all. So it was in general. The more prestigious occupations—law, engineering, science, public relations, and so on—seem to require a certain minimum I.Q., well above the minimum for the less prestigious occupations—for the bakers, chauffeurs, barbers. As far as I.Q. alone is concerned, virtually anyone can be, for example, a welder, but half of mankind (the half below I.Q. 100) is not eligible for auditing, even if the brightest welder may equal the brightest auditor in I.Q.

In this characteristic way, then, I.Q. affects one's occupation. And it is obvious that occupation affects one's social standing. It then follows logically that I.Q. affects social standing. When people are asked to rate the prestige of different occupations, they turn up with lists that look very much like the lists based on average I.Q.'s—the professionals at the top, the laborers at the bottom, and the minor businessmen and white-collar workers in the middle. These ratings have been as stable as the corresponding data on the I.Q., in both America and Europe, and according to people up and down the social scale.

The ties among I.Q., occupation, and social standing make practical sense. The intellectual demands of engi-

neering, for example, exceed those of ditch digging. Hence engineers are brighter, on the average. If virtually anyone is smart enough to be a ditch digger, and only half the people are smart enough to be engineers, then society is, in effect, husbanding its intellectual resources by holding engineers in greater esteem and paying them more. The critics of testing say that the correlations between I.Q. and social class show that the I.Q. test is contaminated by the arbitrary values of our culture, giving unfair advantage to those who hold them. But it is probably no mere coincidence that those values often put the bright people in the prestigious jobs. By doing so, society expresses its recognition, however imprecise, of the importance and scarcity of intellectual ability.

Binet's first scale served the humanitarian goal of getting retarded children into schools that would help them. But the test can also be used to spot exceptionally gifted children, for their own sake and society's. If the tests work, then gifted children should grow up to become unusually accomplished adults, just as the reverse is true at the other end of the scale.

The top of the scale provided the subject of a massive longitudinal study by Lewis M. Terman and his associates at Stanford University. For almost forty years, they followed the lives of a large group of gifted people, publishing their results in five volumes between 1925 and 1959 under the general title of *Genetic Studies of Genius*. The plan of the study was simple: find a large group of young children with exceptionally high I.Q.'s, record as many potentially interesting and useful additional facts about them as practicable, and then follow the course of their lives. Terman and his staff found slightly more than 1500 California children whose I.Q.'s averaged about 150. (Because they used different intelligence scales for some of the children, no precise average figure can be given.) This was no small achievement in itself, for an I.Q. of 150 or greater is, a rarity, possessed, on the average, by the smartest child in a randomly selected group of about two hundred. Most of the children were between the ages of eight and twelve when chosen, but there were also some younger and some recruited in high schools.

Right from the start the findings were informative. For example, highly bright boys were easier to locate than highly bright girls. And the disparity increased slightly with age, suggesting that whatever the I.Q. is, boys maintain it better than girls. For this reason, the final sample had 857 boys and 671 girls. The children, mainly from urban public schools, definitely did not represent the ethnic or social composition of their communities. Compared to the population from which they were drawn, there was an enormous (over tenfold) excess of the children of fathers in the professions and an even more marked scarcity (only .013) of the children of laborers, echoing once again the correlation between I.Q. and social class. In addition, the sample contained an excess of Western and Northern Eu-

If very high income is your goal, and you have a high I.Q., do not waste your time with formal education beyond high school.

Europeans and Jews, and a shortage of Latins, non-Jewish Eastern Europeans, and Negroes. Since the communities sampled had relatively few Orientals, it was hard to tell whether too few, too many, or just the right number of gifted Oriental children turned up, statistically speaking.

The children were non-representative physically as well as intellectually, ethnically, and socially. They tended to be taller, heavier, more broad-shouldered, stronger in hand grip, larger in the vital capacity of their lungs, and somewhat earlier in their sexual maturity than children in the general population. The physical differences, though not large, were large enough to counter the stereotype of the fragile bookworm. Not surprisingly, the gifted children did better in school than their classmates, but mainly in subjects—like reading and arithmetic—that seem to call on intelligence. In subjects like woodworking or sewing, the gifted children enjoyed no particular advantage. They most often liked precisely the subjects that the other children most often disliked, such as reading and arithmetic. At seven years of age the gifted children were already reading books at a higher rate than the average child of fifteen. And even in sports they outdid their classmates, knowing more about the games of childhood and knowing about them earlier. Finally, even in tests of “character”—honesty, tendency toward overstatement, trustworthiness, and the like—the gifted children showed their precocity. At nine or ten years, they had reached the “moral development,” by those no doubt quaint standards, of the average child of thirteen or fourteen.

Children with I.Q.'s of 150 or so are, then, special. But the big question is whether they mature into something special, for that would be the proper test of intelligence testing. Did the I.Q. make the difference it should have made? At last assessment, the sample had reached their middle forties, about thirty-five years after their selection for the study. The death rate in the sample had been less by a third than that in the general population, with fatal accidents quite uncommon. Childhood delinquency, criminal convictions, and alcoholism are all strikingly rare in the sample. More common, and benign, maladjustments are not so rare, with the women showing slightly more emotional trouble than the men. It may be a psychological burden to be so bright a woman in our culture, but this is pure speculation. In any event, not much can be made of the differences in minor mental disturbance between the sample and the general population.

About 70 percent of the sample *finished* college, men ahead of women by a couple of percentage points. This should be compared with the 8 percent of their contemporaries in the general population who finished college (the 1930-1940 college generation). Out of the more than 1500 in the sample, only eleven did not finish high school, and of these, eight went to professional or trade school. Forty percent of the male college graduates earned law, medical, or Ph.D. degrees, and over half of all the college graduates

have at least some postgraduate training. There are, proportionately, five times as many Ph.D.'s in the sample as in the population of college graduates in general. As expected, the sample excelled in college: 80 percent averaging B or better in their courses, and more than 35 percent graduating with honors (Phi Beta Kappa, *cum laude*, or the like). In addition to their academic degrees, the sample has earned a disproportionately large number of professional licenses—CPA's, Fellows of the American Board of Surgery, Fellows of the American Institute of Architects, and so on.

The ten most common occupations among the men are not the common lot in our society: lawyers first, followed by college faculty members, engineers, physicians, school administrators or teachers, chemists and physicists, authors, architects, geologists, and clergymen. All told, over 85 percent of the working men became either professionals or managers in business and industry, with the first category the larger. At the other end of the occupational scale, only about 3 percent became semiskilled laborers or farmers, and virtually none unskilled laborers. The men are bunched at the top of the scale of occupations, just as they are at the top of the scale of I.Q. And the sample outperforms not only the population in general, but also the average college graduate. The run-of-the-mill college graduate has a 5 percent chance of becoming a semiskilled or unskilled laborer; the sample's college graduate has a chance of only .5 percent, a tenfold reduction.

Even with the relatively fewer employed women than men in the sample, the distinction of employment still shows. Approximately two thirds of the working women held professional positions—in universities, school and welfare systems, journalism, medical and paramedical professions, and so on. In addition, most of the women were married and raising children. The antifeminist threat that the education of women could remove the brightest potential mothers from the breeding stock receives no support in these results.

In addition to everything else, a high I.Q. pays in money. The average professional or managerial man in the sample was earning about \$10,500 in 1954, compared to a national average of about \$6000 for those occupations. Even the semiskilled and clerical workers in the sample were outearning, by about 25 percent, the general averages for the same jobs. The total family income for the sample more than doubled that for white, urban American families of roughly the same socio-economic status. About 30 percent of the families in the sample earned more than \$15,000 a year in 1954, compared to only one percent for ordinary families in the same general socio-economic class. The sample shows the economic advantages of a high I.Q., after discounting education, race, occupation, and geography.

In the general population, income and education correlate highly. One encounters, from time to time, estimates

of how much a high school diploma or a bachelor's degree should add to one's paycheck. No doubt about the facts—more highly educated people make more money—but the interpretation is arguable. The usual interpretation assumes that a given man with some higher education would earn more than the same man *without* the education. But that is not really what the data show, for we do not know whether the people who have the extra education are the same in other ways as those who do not. Suppose, for the sake of relevance, that income *really* depended more on I.Q. than on education. Suppose further that the amount of schooling also depended on I.Q. Educated people would then earn more not because they were more educated (which they would also be), but because they were smarter (had a higher I.Q.). To disentangle the complex factors in society at large and find out what causes what, simple correlations are not enough. We need to know if income would be correlated with I.Q. if education were held constant, or conversely if income would be correlated with education if I.Q. were held constant. Terman's high I.Q. sample is a step in this direction, for it allows us to see whether income depends on education when I.Q. is held constant at the virtual top of the scale.

High-school graduates in the Terman sample were earning about as much as the college graduates with a bachelor's degree. Further schooling beyond the bachelor's did improve income somewhat. However, of the six men with the highest incomes (ranging upward of \$100,000 a year) only one finished college. The highest annual income of all, \$400,000 for the last year reported, was earned by a man who had had no college whatever. In other words, if *very* high income is your goal, and you have a high I.Q., do not waste your time with formal education beyond high school. For this particular sample, education did not unequivocally add to income, as most often claimed. Perhaps for people of more ordinary talents the connection is straightforward, but the study shows that while I.Q. definitely affects income, education may not.

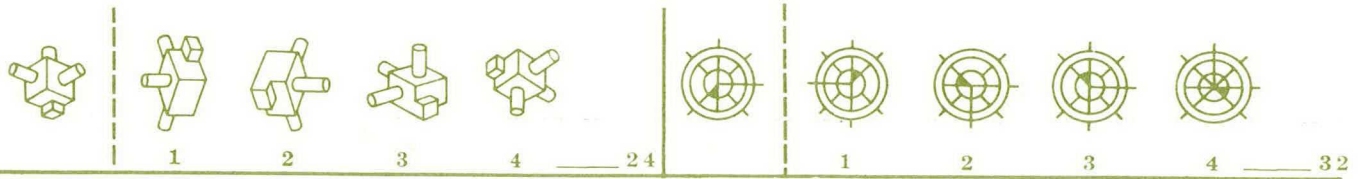
Not just the economic facts of life were gathered. When the men were asked about their state of mind, almost 90 percent said that they were at least fairly content, and virtually half were finding "deep satisfaction" in their lives. Only 6 percent reported discontentment. The more prosperous men were generally the more contented. When the men were asked to estimate how well they were living up to their intellectual abilities, there was again a correlation between satisfaction and income. The average yearly wage of those who said they were "fully" living up to their capacities was almost \$12,000, while the group least satisfied was making less than \$5000 per year.

Women's salaries were substantially lower than the men's and did not correlate with contentment. Notwithstanding their poorer salaries, on an average, the women reported greater satisfaction in their lives than the men. The housewives, who were earning less money than

anyone else, expressed about as much satisfaction as any other group in the sample. There is little here to support the feminist argument that a housewife's life is intolerable, especially for educated, intelligent women. It would be hard to pick a brighter group than the women in this study, yet they seemed to be adjusting easily to their lot. To give but one example of the many striking cases, a woman whose I.Q. of 192 places her close to the top of the entire sample, and whose retested intelligence at maturity was again virtually at the top, was raising eight children, including three sets of twins. According to the account at the latest report, she had no outside activity at that time other than an interest in the P.T.A., but was apparently content, if not serene. Of course, such tranquillity may be gone now, fifteen years later.

The enormous harvest from the sample in their middle forties included about 2000 scientific and technical articles, 60 books, 33 novels, 375 short stories or plays, 325 miscellaneous publications, 230 patents, not to mention the hundreds of radio and television scripts, newspaper stories, pieces of art and music. Their names turn up disproportionately often in compilations of our effective people—in *Who's Who*, *American Men of Science*, the National Academy of Sciences, and so on. They are active in P.T.A., clubs, hobbies. They vote far more faithfully (over 90 percent of the time in national elections) than the general population (and are somewhat more conservative). By the mid-1950s, they had spawned (with spouses who were themselves significantly brighter than average) about 2500 children whose average I.Q. appears to be above 130—not as brilliant as their exceptional parents, but still among the top 5 percent of the population. Even in their mid-forties, the sample continued to test within the top one percent of the general population in intelligence, whether or not they had been successful in their careers. No doubt the predictive power of the I.Q. is outlasting the first thirty-five years of the study.

No single study is beyond criticism, not even this massive enterprise by Terman and his associates. Critics can point to the possibility of hidden biases in the original selection of the children. Not *every* child in the California schools was tested, only those who looked "promising" for one reason or another. The final selection employed just the I.Q., but the pre-screening may indeed have been a source of bias. Later estimates uncovered, however, only a few children missed this way, certainly not enough to change the general conclusions about the predictiveness of I.Q. Critics may also wonder how people are affected by being included in this select group. Are they impelled to excel, or are they stunted by anxiety? Judging from all the other data showing correlations between I.Q. and achievement, the sample seems to be about normal for an I.Q. of 150. Whatever the flaws in the study, there can be no reasonable doubt about its main conclusion. An I.Q. test can be given in an hour or two to a child, and from this in-



finitesimally small sample of his output, deeply important predictions follow—about schoolwork, occupation, income, satisfaction with life, and even life expectancy. The predictions are not perfect, for other factors always enter in, but no other single factor matters as much in as many spheres of life.

Terman was unapologetic about where he thought I.Q. comes from. He believed in the inheritance of I.Q., at least to a considerable degree. Bluntly, but not dogmatically, he wrote in 1925:

There are . . . many persons who believe that intelligence quotients can be manufactured to order by the application of suitable methods of training. There are even prominent educators and psychologists who are inclined to regard such a pedagogical feat as within the realm of possibility, and no one knows that it is not. If it is possible it is time we were finding out. Conclusive evidence as to the extent to which I.Q.'s can be artificially raised could be supplied in a few years by an experiment which would cost a few hundred thousand or at most a few million dollars. The knowledge would probably be worth to humanity a thousand times that amount.

The opening paragraphs of the disturbing and controversial article by Professor Arthur R. Jensen of the University of California, Berkeley, could be taken as the equally blunt answer to Terman's challenge, forty-four years later.

Compensatory education has been tried and it apparently has failed.

Compensatory education has been practiced on a massive scale for several years in many cities across the nation. It began with auspicious enthusiasm and high hopes of educators. It had unprecedented support from Federal funds. It had theoretical sanction from social scientists espousing the major underpinning of its rationale: the "deprivation hypothesis," according to which academic lag is mainly the result of social, economic, and educational deprivation and discrimination—an hypothesis that has met with wide, uncritical acceptance in the atmosphere of society's growing concern about the plight of minority groups and the economically disadvantaged.

The chief goal of compensatory education—to remedy the educational lag of disadvantaged children and thereby narrow the achievement gap between "minority" and "majority" pupils—has been utterly unrealized in any of the large compensatory education programs that have been evaluated so far.

And the reason, Jensen goes on to say, why compensatory education has failed is that it has tried to raise I.Q.'s, which, he argues, are more a matter of inheritance

than environment, and therefore not very amenable to corrective training. What evidence has he for this unexpected and unpopular conclusion?

The problem with nature and nurture is to decide which—inheriting or environment—is primary, for the I.Q. is exclusively the result of neither one alone. Advocates of environment—the clear majority of those who express themselves publicly on the subject—must explain why I.Q.'s usually stay about the same during most people's lives and also why high or low I.Q.'s tend to run in families. Those facts could easily be construed as signs of a genetic basis for the I.Q. The usual environmentalist answer argues that I.Q.'s remain the same to the extent that environments remain the same. If you are lucky enough to be wellborn, then your I.Q. will show the benefits of nurturing, which, in turn, gives you an advantage in the competition for success. If, on the other hand, you are blighted with poor surroundings, your mental growth will be stunted and you are likely to be stuck at the bottom of the social ladder. By this view, parents bequeath to their children not so much the genes for intelligence as the environment that will promote or retard it.

In one plausible stroke the environmentalist arguments seem to explain, therefore, not only the stability of the I.Q. but also the similarity between parents and children. The case is further strengthened by arguing that early training fixes the I.Q. more firmly than anything we know how to do later. And then to cap it off, the environmentalist may claim that the arbitrary social barriers in our society trap the underprivileged in their surroundings while guarding the overprivileged in theirs. Anyone who accepts this series of arguments is unshaken by Jensen's reminder that compensatory education has failed in the United States, for the answer seems to be ready and waiting. To someone who believes in the environmental theory, the failure of compensatory education is not disproof of his theory, but rather a sign that we need more and better special training earlier in a person's life.

To be sure, it seems obvious that poor and unattractive surroundings will stunt a child's mental growth. To question it seems callous. But even if it is plausible, how do we know it is true? By what evidence do we test the environmentalist doctrine? The simplest possible assessment of the inherited factor in I.Q. is with identical twins, for only environmental differences can turn up between people with identical genes. In an article recently published in the periodical *Behavior Genetics*, Professor Jensen surveys four major studies of identical twins who were reared in separate homes. Most of the twins had been separated by the age of six months, and almost all by the age of two years. The twins were Caucasians, living in England, Denmark, and the United States—all told, 122 pairs

“How Much Can We Boost I.Q. and Scholastic Achievement?” “Not very much.”

of them. The overall I.Q. of the 244 individuals was about 97, slightly lower than the standard 100. Identical twins tend to have slightly depressed I.Q.'s, perhaps owing to the prenatal hazards of twindom. The 244 individuals spanned the range of I.Q.'s from 63 to 132, a range that brackets most of humanity—or to be more precise, 97 percent of the general population on whom intelligence tests have been standardized.

Being identical twins, the pairs shared identical genetic endowments, but their environments could have been as different as those of random pairs of children in the society at large. Nevertheless, their I.Q.'s correlated by about 85 percent, which is more than usual between ordinary siblings or even fraternal twins growing up together with their own families. It is, in fact, almost as big as the correlations between the heights and weights of these twins, which were 94 percent and 88 percent respectively. Even environmentalists would expect separately raised twins to look alike, but these results show that the I.Q.'s match almost as well. Of course if the environment alone set the I.Q., the correlations should have been much smaller than 85 percent. It would, however, be rash to leap to the conclusion that the 85 percent correlation is purely genetic, for when twins are placed into separate homes, they might well be placed into similar environments. The children had been separated not for the edification of psychologists studying the I.Q., but for the weighty reasons that break families up—illness, poverty, death, parental incapacity, and so on—and the accidents of separation may not have yielded well-designed experiments. Some of the pairs were no doubt raised by different branches of the same family, perhaps assuring them considerable environmental similarity anyway. In such cases, the correlation of 85 percent would not be purely genetic, but at least partly environmental. Fortunately for our state of knowledge, one of the four studies examined by Jensen included ratings of the foster homes in terms of the breadwinner's occupation. Six categories sufficed: higher professional, lower professional, clerical, skilled, semiskilled, unskilled. Now, with this classification of homes, we know a little about whether the twins were raised in homes with a similar cultural ambience. To the extent that the environment in a home reflects the breadwinner's occupation, the answer is unequivocally negative, for there was literally no general correlation in the occupational levels of the homes into which the pairs were separated. At least for this one study—which happened to be the largest of the four—the high correlation in I.Q. resulted from something besides a social-class correlation in the foster homes, most likely the shared inheritance.

Twins raised apart differ on the average by about seven points in I.Q. Two people chosen at random from the general population differ by seventeen points. Only four of the 122 pairs of twins differed by as much as seventeen points. Ordinary siblings raised in the same household dif-

fer by twelve points. Only nineteen of the 122 twin pairs differed by as much as that. And finally, fraternal twins raised in the same home differ by an average of eleven points, which was equaled or exceeded by only twenty-three of the 122 pairs. In other words, more than four times out of five the difference between identical twins raised apart fell short of the average difference between fraternal twins raised together by their own parents. At the same time, those separated twins were not so similar in schoolwork. Identical twins raised together resemble each other in both I.Q. and school grades. When twins are separated, their I.Q.'s remain quite close, but their grades diverge. It seems that school performance responds to the environment substantially more than does the I.Q., although neither one is solely the outcome of either nature or nurture.

The comparison between I.Q. and grades was one theme of Jensen's controversial earlier article, “How Much Can We Boost I.Q. and Scholastic Achievement?”, which appeared in the winter of 1969 in the *Harvard Educational Review*. Jensen answered the title's rhetorical question about I.Q. with a scholarly and circumspect form of “not very much.” The article is cautious and detailed, far from extreme in position or tone. Not only its facts but even most of its conclusions are familiar to experts. The failure of compensatory education was the occasion for the article, which served especially well in assembling many scattered but pertinent items. Jensen echoes most experts on the subject of the I.Q. by concluding that substantially more can be ascribed to inheritance than environment. Since the importance of inheritance seems to say something about racial differences in I.Q. that most well-disposed people do not want to hear, it has been argued that Jensen should not have written on the subject at all or that the *Harvard Educational Review* should not have, as it did, invited him to write on it.

Some of Jensen's critics have argued that because environment and inheritance are intertwined, it is impossible to tease them apart. The criticism may seem persuasive to laymen, for nature and nurture are indeed intertwined, and in just the way that makes teasing them apart most difficult. For intelligence—unlike, for example, skin color—the main agents of both nature and nurture are likely to be one's parents. One inherits skin color from one's parents, but the relevant environment does not come directly from them but from sun, wind, age, and so on. For skin color, resemblance to parents signifies (albeit not infallibly) inheritance; for intelligence, resemblance is ambiguous. Nevertheless analysis is possible even with I.Q., as Jensen and his predecessors have shown. The most useful data for the purpose are the correlations between I.Q. and kinship, as exemplified by the twin studies, which set genetic similarity high and environmental similarity low. Foster children in the same home define the other extreme of kinship and environment. If environment had no bearing at all on

Jensen concluded (as have most of the other experts in the field) that the genetic factor is worth about 80 percent and that only 20 percent is left to everything else.

intelligence, then the I.Q.'s of such unrelated children should correlate slightly at most (and only to the extent caused by a special factor to be mentioned shortly). In contrast, if environment were all, then the correlation should approach the value for natural siblings. Actually, the I.Q.'s of foster children in the same home correlate by about 24 percent (less than half the value for natural siblings). However, even the correlation of 24 percent cannot be credited entirely to the children's shared environment. Bear in mind that adoption agencies try to place "comparable" children in the same home, which means that there is more than just their common surroundings making them alike. Suppose, for example, that adoption agencies tried to put children with similar hair color in any given family. They could check on the natural parents, and perhaps even on the grandparents, and make a reasonable guess about the baby's eventual hair color. The foster children in a given home would then often have similar hair color; they would be unrelated by blood, but the similarity would be more genetic than environmental. By trying for a congenial match between foster child and foster parents—in appearance and in mental ability—adoption agencies make the role of environment look more important than it probably is.

In between foster siblings and identical twins come the more familiar relations, and these too have been scrutinized. If intelligence were purely genetic, the I.Q.'s of second cousins would correlate by 14 percent and that of first cousins by 18 percent (the reasons for those peculiar percentages are well beyond the scope of this article, so they are offered without proof). Instead of 14 percent and 18 percent, the actual correlations are 16 percent and 26 percent—too large for genetic influences alone, but in the right range. Uncle's (or aunt's) I.Q. should, by the genes alone, correlate with nephew's (or niece's) by a value of 31 percent; the actual value is 34 percent. The correlation between grandparent and grandchild should, on genetic grounds alone, also be 31 percent, whereas the actual correlation is 27 percent, again a small discrepancy. And finally for this brief survey, the predicted correlation between parent and child, by genes alone, is 49 percent, whereas the actual correlation is 50 percent using the parents' adult I.Q.'s and 56 percent using the parents' childhood I.Q.'s—in either case too small a difference to quibble about. Parents and their children correlate about as well whether the children are raised at home or by a foster family, which underscores the relative unimportance of the environment.

The foregoing figures are lifted directly out of Jensen's famous article, figures that he himself culled from the literature of intelligence testing. The measurements say that (1) the more closely related by blood two people are, the greater the correlation between their I.Q.'s and (2) the correlations fall in the right range from the purely genetic standpoint. By evaluating the total evidence, and by a

procedure too technical to explain here, Jensen concluded (as have most of the other experts in the field) that the genetic factor is worth about 80 percent and that only 20 percent is left to everything else—the social, cultural, and physical environment, plus illness, prenatal factors, and what have you.

Jensen's two papers leave little doubt about the heritability of I.Q. among North American and Western European whites, whom most data on the subject describe. In fact, there is little dispute on this score, even among those who object vigorously to this work. It is the relation between heritability and racial differences that raises the hackles. Given the well-established, roughly fifteen-point black-white difference in I.Q., the argument is whether the difference arises in the environment or the genes. If intelligence were entirely genetic, then racial differences would be genetic simply because they could be due to nothing else. Conversely, if the genes were irrelevant, then the racial difference would have to be due to the environment, again because there would be no alternative. As it is, I.Q. reflects both a person's genes and his environment. The racial issue really poses the nature-nurture question all over again, but this time for a particular finding—the higher scores of whites over blacks on I.Q. tests.

In general—not just for the racial issue—the question of nature and nurture boils down to the study of variation. Granted that I.Q.'s vary among people, to what extent does the variation correlate with the differences in their surroundings on the one hand and with the differences in their genetic makeup on the other? No one disputes the existence of all three kinds of variation—in I.Q., environment, and inheritance—only their interconnections. In effect, the environmentalist is saying that among a group of people, the various I.Q.'s reflect the various surroundings more or less without regard to the genes. In contrast, the nativist is saying the reverse—that different I.Q.'s reflect different genetic endowments rather than different environments. The study of quantitative genetics contrives to answer such riddles, and so a brief didactic excursion is in order. But instead of starting the lesson with I.Q., let us consider a trait which we are not emotionally committed to to begin with.

Suppose we wanted to know the heritability of skin color. We would not need science to tell us that dark or fair complexions run in certain families or larger groups. Nor must we be told that nongenetic elements also enter in, as when a person is tan from the sun or pale with illness or yellow from jaundice or red with rage or blue with cold. The task of quantitative genetics is to come up with a number that says how large a role inheritance plays in the total amount of variation in skin color that we see in a particular group of people at a particular time. If the number is large, then skin color is largely heritable; if very small,

then the heritability is negligible. If the number is large, then there will be marked family resemblances; if small, then members of given families will be no more alike than unrelated people. To convey such information, the number must reflect which group of people we choose to study. Consider first the United States, with its racial and ethnic diversity. Much skin variation here is related to ancestry, whether black, white, yellow, red, or Mediterranean, Nordic, Alpine, or some blend. Family resemblances in skin color are quite strong in America, so the heritability should come out large. Now contrast this with an isolated village in Norway, full of Scandinavians with generations of pale-skinned ancestors. In the Norwegian town, whatever little variation there is in skin color is likely to be environmental, due to the circumstances of life rather than to the accident of inheritance. As regards skin color, children will be no more like their parents than their nonrelatives, so heritability should come out low.

The hardest thing to grasp about heritability is that it says something about a trait in a population as a whole, not about the relation between particular parents and their offspring. Skin color turns out to be more heritable in the United States than in Norway, even though the physiological mechanisms of inheritance are surely the same. In the Norwegian town, a swarthy father and mother (who probably got that way from exposure to the weather) are likely to have children as fair-skinned as their neighbors. In the American town, however, it is more likely that the swarthy parents is genetic and will be passed on to the children. Although heritability is not the strictly physiological concept that laymen imagine it to be, it is uniquely useful for talking about the nature-nurture question, for it tells us whether traits run in families within a broader population of individuals.

The technical measure of heritability is a number between 0 and 1.0 that states how much of the variation in a trait is due to genetic factors. How it is calculated need not detain us here. It is enough to know that a heritability of .5 means (omitting some technical complexities) that the variation is due half to genetic factors and half to other factors; a heritability of .2 means that only a fifth of the variation is genetic, and so on. Some actual heritabilities of traits in animals may be helpful. In piebald Holstein cattle, for example, the amount of white in the fur has a heritability of about .95, a value so high that it is almost right to say the environment plays no role here. In contrast, milk yield has a heritability of only .3. White in the fur, therefore, breeds more true than milk production. In pigs, the thickness of body fat has a heritability of .55, while the litter size has a heritability of only .15.

Now back to I.Q. and the racial issue. Using the procedures of quantitative genetics, Jensen (and most other experts) estimates that I.Q. has a heritability between .80 and .85, but this is based almost entirely on data from whites. We may, therefore, say that 80 to 85 percent of the vari-

ation in I.Q. among whites is due to the genes. Because we do not know the heritability for I.Q. among blacks, we cannot make a comparable statement for them. But let us simply assume, for the sake of discussion, that .8 is the heritability for whites and blacks taken together. What could we say about the racial difference in I.Q. then? The answer is that we could still say nothing positive about it. Recall that the concept of heritability applies to a population as a whole. All we could say is that the differences between people, on the average and without regard to color, are 80 percent inherited. But within this broad generality, particular differences could and would be more or less inherited. Take, for example, the differences in I.Q. between identical twins. Even with the average heritability equal to .8, all twin differences have to be totally environmental, since their genes cannot differ. Or conversely, consider the differences between foster children in a given foster family. Because they are growing up in the same home, their I.Q. differences could easily be relatively more genetic than those of people taken at random. When this line of reasoning is applied to a racial (or ethnic) difference in I.Q., the only proper conclusion is that we do not know whether it is more genetic, less genetic, or precisely as genetic as implied by a heritability of .8.

Jensen notes that we lack a good estimate of the heritability of intelligence among blacks. Although there are scraps of evidence for a genetic component in the black-white difference, the overwhelming case is for believing that American blacks have been at an environmental disadvantage. To the extent that variations in the American social environment can promote or retard I.Q., blacks have probably been held back. But a neutral commentator (a rarity these days) would have to say that the case is simply not settled, given our present stage of knowledge. To advance this knowledge would not be easy, but it could certainly be done with sufficient ingenuity and hard work. To anyone who is curious about the question and who feels competent to try to answer it, it is at least irritating to be told that the answer is either unknowable or better not known, and both enjoinders are often heard. And there is, of course, a still more fundamental issue at stake, which should concern even those who are neither curious about nor competent to study racial differences in I.Q. It is whether inquiry shall (again) be shut off because someone thinks society is best left in ignorance.

Setting aside the racial issue, the conclusion about intelligence is that, like other important though not necessarily vital traits, it is highly heritable. It is not vital in the sense that it may vary broadly without markedly affecting survival, although it no doubt affects one's life-style. Does it do us any practical good to know how heritable intelligence is? We are not, for example, on the verge of Galton's vision of eugenics, even though we now have the mental test that he thought was the crucial prerequisite. For good or ill, and for some time to come, we are stuck

The “regression toward the mean” is the tendency for children to be closer to the general population average (in this case, I.Q. 100) than their parents. Very bright parents have children who tend to be merely bright, while very dull parents tend to have them merely dull.

with mating patterns as people determine them for themselves. No sensible person would want to entrust state-run human breeding to those who control today's states. There are, however, practical corollaries of this knowledge, more humble than eugenics, but ever more salient as the growing complexity of human society makes acute the shortage of high-grade intellect.

Heritability is first and foremost the measure of breeding true, useful for predicting how much of some trait the average offspring in a given family will have. For example, to predict the I.Q. of the average offspring in a family:

1. Average the parents' I.Q.'s.
2. Subtract 100 from the result.
3. Multiply the result of (2) by .8 (the heritability).
4. Add the result of (3) to 100.

Thus, given a mother and father each with I.Q.'s of 120, their average child will have an I.Q. of 116. Some of their children will be brighter and some duller, but the larger the family, the more nearly will the average converge onto 116. With parents averaging an I.Q. of 80, the average child will have an I.Q. of 84. The formula predicts something the experts call “regression toward the mean,” the tendency for children to be closer to the general population average (in this case, I.Q. 100) than their parents. And in fact, *very bright* parents have children who tend to be merely *bright*, while *very dull* parents tend to have them merely *dull*. The amount of regression for a trait depends on the heritability—with high heritability, the regression is smaller than with low. Also, for a given trait the regression is greater at the extremes of a population than at its center. In other words, ordinary parents are more like their children (on the average) than extraordinary ones (whether extraordinarily high or low). All of these characteristics of the “generation gap” follow directly and completely from the simple formula given above. Thus, when the parents average 120, the regression effect is only four I.Q. points, but if they averaged 150, the regression effect would be ten points. In comparison, height, with its heritability of .95, would show smaller regression effects than I.Q., since the multiplier in step 3 of the formula is closer to 1.0. But even so, very tall parents tend to have children who are merely tall, and very short parents tend to have them merely short. As long as the heritability of a trait falls short of 1.0, there is some regression effect.

Intelligence may be drifting up or down for environmental reasons from generation to generation, notwithstanding the high heritability. Height, for example, is said to be increasing—presumably because of diet and medicine—even with its .95 heritability. We can easily tell whether there has been a change in height, for the measures are absolute, and there is the tangible evidence of clothing, furniture, coffins, and the skeletons themselves. For intelligence, however, we have no absolute scales, only relative ones, and the tangible remains of intelligence defy interpretation. But if height has changed, why not in-

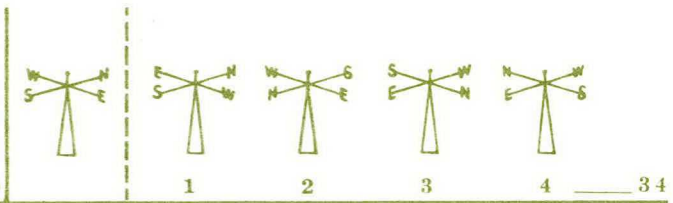
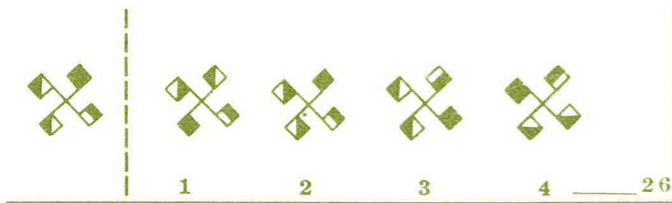
telligence? After all, one could argue, the I.Q. has a heritability of only .8, measurably lower than that of height, so it should be even more amenable to the influence of the environment. That, to be sure, is correct in principle, but the practical problem is to find the right things in the environment to change—the things that will nourish the intellect as well as diet does height. The usual assumption, that education and culture are crucial, is running into evidence that the physical environment—for example, early diet—might be more important. In fact, the twin studies that Jensen surveyed showed that the single most important environmental influence on I.Q. was not education or social environment, but something prenatal, as shown by the fact that the twin heavier at birth usually grew up with the higher I.Q.

Suppose we do find an environmental handle on I.Q.—something, let us say, in the gestating mother's diet. What then? Presumably society would try to give everyone access to the favorable factor, within the limits of its resources. Intelligence would increase accordingly. But that would not end our troubles with I.Q. Recall that heritability is a measure of relative variation. Right now, about 80 percent of the variation in I.Q. derives from the genes. If we make the relevant environment much more uniform (by making it as good as we can for everyone), then an even larger proportion of the variation in I.Q. will be attributable to the genes. The average person would be smarter, but intelligence would run in families even more obviously and with less regression toward the mean than we see today. It is likely that the mere fact of heritability in I.Q. is socially and politically important, and the more so the higher the heritability.

The specter of Communism was haunting Europe, said Karl Marx and Friedrich Engels in 1848. They could point to the rise of egalitarianism for proof. From Jefferson's “self-evident truth” of man's equality, to France's “*égalité*” and beyond that to the revolutions that swept Europe as Marx and Engels were proclaiming their *Manifesto*, the central political fact of their times, and ours, has been the rejection of aristocracies and privileged classes, of special rights for “special” people. The vision of a classless society was the keystone of the Declaration of Independence as well as the *Communist Manifesto*, however different the plan for achieving it.

Against this background, the main significance of intelligence testing is what it says about a society built around human inequalities. The message is so clear that it can be made in the form of a syllogism:

1. If differences in mental abilities are inherited, and
2. If success requires those abilities, and
3. If earnings and prestige depend on success,
4. Then social standing (which reflects earnings and



prestige) will be based to some extent on inherited differences among people.

The syllogism has five corollaries, which make it more relevant to the future than to the past or present.

a) As the environment becomes more favorable for the development of intelligence, its heritability will increase, as the preceding section showed. Regardless of whether this is done by improving educational methods, diet for pregnant women, or whatever, the more advantageous we make the circumstances of life, the more certainly will intellectual differences be inherited. And the greater the heritability, the greater the force of the syllogism.

b) All modern political credos preach social mobility. The good society should, we believe, allow people to rise (and, by implication if not by frank admission, fall) according to their own efforts. The social barriers of the past—race, religion, nationality, title, inherited wealth—are under continuous assault, at least in principle. The separation of church and state, the graduated income tax, the confiscatory inheritance tax, the laws against discrimination and segregation, the abolition of legal class and caste systems all manifest a desire to accelerate movement on the social ladder. The standard wisdom of our time avows that people should be free of “unfair” impediments and divested of “unfair” advantages in all their endeavors. But the syllogism becomes more potent in proportion to the opportunities for social mobility, for it is only when able and energetic individuals can rise and displace the dull and sluggish ones that there can be sorting out of people according to inherited differences. Actual social mobility is blocked by innate human differences after the social and legal impediments are removed.

c) It was noted earlier that there are many bright but poor people even in affluent America. The social ladder is tapered steeply, with far less room at the top than at the bottom. The obvious way to rescue the people at the bottom is to take the taper out of the ladder, which is to say, to increase the aggregate wealth of society so that there is more room at the top. This is, of course, just what has been happening since the Industrial Revolution. But one rarely noted by-product of poverty is that it minimizes the inherited differences between classes by assuring that some bright people will remain at the bottom of the ladder. As the syllogism implies, when a country gains new wealth, it will tend to be gathered in the hands of the natively endowed. In other words, the growth of wealth will recruit for the upper classes precisely those from the lower classes who have the edge in native ability. Whatever else this accomplishes, it will also increase the I.Q. gap between upper and lower classes, making the social ladder even steeper for those left at the bottom.

d) Technological advance changes the marketplace for I.Q. Even if every single job lost in automating a factory is replaced by a new job someplace else in a new technology, it is more than likely that some of those put out of the old

jobs will not have the I.Q. for the new ones. Technological unemployment is not just a matter of “dislocation” or “re-training” if the jobs created are beyond the native capacity of the newly unemployed. It is much easier to replace men’s muscles with machines than to replace their intellects. The computer visionaries believe that their machines will soon be doing our thinking for us too, but in the meantime, backhoes are putting ditchdiggers out of work. And the ones who stay out of work are most likely the ones with the low I.Q.’s. The syllogism implies that in times to come, as technology advances, the tendency to be unemployed may run in the genes of a family about as certainly as bad teeth do now.

e) The syllogism deals manifestly with intelligence. The invention of the intelligence test made it possible to gather the data necessary to back up the three premises. However, there may be other inherited traits that differ among people and contribute to their success in life. Such qualities as temperament, personality, appearance, perhaps even physical strength or endurance, may enter into our strivings for achievement and are to varying degrees inherited. The meritocracy concerns not just inherited intelligence, but all inherited traits affecting success, whether or not we know of their importance or have tests to gauge them.

The syllogism and its corollaries point to a future in which social classes not only continue but become ever more solidly built on inborn differences. As the wealth and complexity of human society grow, there will be precipitated out of the mass of humanity a low-capacity (intellectual and otherwise) residue that may be unable to master the common occupations, cannot compete for success and achievement, and are most likely to be born to parents who have similarly failed. In Aldous Huxley’s *Brave New World*, it was malevolent or misguided science that created the “alphas,” “gammas,” and the other distinct types of people. But nature itself is more likely to do the job or something similar, as the less well-known but far more prescient book by Michael Young, *The Rise of the Meritocracy*, has depicted. Young’s social-science-fiction tale of the antimercitocratic upheavals of the early twenty-first century is the perfect setting for his timely neologism, the word “meritocracy.” The troubles he anticipated, and that the syllogism explains, have already caught the attention of alert social scientists, like Edward Banfield, whose book *The Unheavenly City* describes the increasingly chronic lower class in America’s central cities. While Sunday supplements and popular magazines crank out horror stories about genetic engineering (often with anxious but self-serving testimonials from geneticists), our society may be sorting itself willy-nilly into inherited castes. What is most troubling about this prospect is that the growth of a virtually hereditary meritocracy will arise out of the successful realization of contemporary political and social goals. The more we succeed in achieving rela-

What is most troubling is that the growth of a virtually hereditary meritocracy will arise out of the successful realization of contemporary political and social goals.

tively unimpeded social mobility, adequate wealth, the end of drudgery, and wholesome environment, the more forcefully does the syllogism apply.

Are there alternatives short of turning back to social rigidity, poverty, drudgery, and squalor? The first two premises of the syllogism cannot sensibly be challenged, for they are true to some extent now and are likely to become more so in time. The heritability of intelligence will grow as the conditions of life are made more uniformly wholesome; intelligence will play an increasingly important role in occupational success as the menial jobs are taken over by machines. The one even plausible hope is to block the third premise by preventing earnings and prestige from depending upon successful achievement. The socialist dictum, "From each according to his ability, to each according to his needs," can be seen as a bald denial of the third premise. It states that, whatever a person's achievement, his income (economic, social, and political) is unaffected by his success. Instead, the dictum implies, people will get what they need however they perform, but only so long as they fulfill their abilities. Those in power soon discover that they must insist on a certain level of performance, for what the dictum neglects is that "ability" is, first of all, widely and innately variable, and secondly, that it expresses itself in labor only for gain. In capitalist countries, the gain is typically in material wealth, but even where the dictum rules (if such places exist), social and political influence or relief from threat would be the reward for accomplishment. Human society has yet to find a working alternative to the carrot and the stick. Meanwhile, the third premise assures the formation of social classes.

Classlessness is elusive because people vary and because they compete for gain—economic and otherwise. The tendency to respect, honor, remunerate, and perhaps even envy people who succeed is not only ingrained but is itself a source of social pressure to contribute to one's limit. Imagine, for example, what would happen if the gradient of gain were inverted by government fiat. Suppose bakers and lumberjacks got the top salaries and the top social approval, while engineers, physicians, lawyers, and business executives got the bottom. Soon thereafter, the scale of I.Q.'s would also invert, with the competition for the newly desirable jobs now including people with the highest I.Q.'s. (For simplicity's sake, only I.Q. is mentioned, but there may be, and no doubt are, other factors that contribute to success, for recall that I.Q. is only necessary, not sufficient.) The top I.Q.'s would once again capture the top of the social ladder. But no government (let alone people themselves) is likely to conduct such an experiment, for it is not a sensible allocation of a scarce resource like high-grade intelligence. Nor could a government long equalize the gains from all occupations. It was noted before that the premium given to lawyers, doctors, engineers, and business managers is not accidental, for those jobs are left to incompetents at our collective peril. There are simply

fewer potentially competent physicians than barbers. The gradient of occupations is, then, a natural measure of value and scarcity. And beneath this gradient is a scale of inborn ability, which is what gives the syllogism its unique potency.

It seems that we are indeed stuck with the conclusion of the syllogism. The data on I.Q. and social-class differences show that we have been living with an inherited stratification of our society for some time. The signs point to more rather than less of it in the future, assuming that we are not plunged back into a state of primeval poverty by some cataclysm or do not turn back to rigidly and arbitrarily privileged classes. Recall that regression toward the mean depends upon the heritability and that improving the environment raises the heritability. The higher the heritability, the closer will human society approach a virtual caste system, with families sustaining their position on the social ladder from generation to generation as parents and children are more nearly alike in their essential features. The opportunity for social mobility across classes assures the biological distinctiveness of each class, for the unusual offspring—whether more or less able than his (or her) closest relatives—would quickly rise above his family or sink below it, and take his place, both biologically and socially, with his peers.

If this is a fair picture of the future, then we should be preparing ourselves for it instead of railing against its dawning signs. Greater wealth, health, freedom, fairness, and educational opportunity are *not* going to give us the egalitarian society of our philosophical heritage. It will instead give us a society sharply graduated, with ever greater innate separation between the top and the bottom, and ever more uniformity within families as far as inherited abilities are concerned. Naturally, we find this vista appalling, for we have been raised to think of social equality as our goal. The vista reminds us of the world we had hoped to leave behind—aristocracies, privileged classes, unfair advantages and disadvantages of birth. But it is different, for the privileged classes of the past were probably not much superior biologically to the downtrodden, which is why revolutions had a fair chance of success. By removing arbitrary barriers between classes, society has encouraged the creation of biological barriers. When people can freely take their natural level in society, the upper classes will, virtually by definition, have greater capacity than the lower.

The measurement of intelligence is one of the yardsticks by which we may assess the growing meritocracy, but other tests of human potential and performance should supplement the I.Q. in describing a person's talents, interests, skills, and shortcomings. The biological stratification of society would surely go on whether we had tests to gauge it or not, but with them a more humane and tolerant grasp of human differences is possible. And at the moment, that seems our best hope. □

Copyright of Atlantic Magazine Archive is the property of Atlantic Monthly Group LLC and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.