The Role of Intelligence in Modern Society

Are social changes dividing us into intellectual haves and have-nots? The question pushed aside in the 1970s is back, and the issues are far from simple

Earl Hunt

Earl Hunt is professor of psychology and adjunct professor of computer science at the University of Washington. He specializes in the study of cognition in people and in artificial-intelligence systems. Within these fields he has focused on individual differences in thinking and their implications for the role of people in a high-technology society. His book, Will We Be Smart Enough?, *to be published in August by the Russell Sage Foundation, combines cognitive theory, demographic projections and psychometric research to measure the capabilities of tomorrow's workforce against the needs of tomorrow's workplace. Address: Department of Psychology, University of Washington, Box 351525, Seattle, WA 98195-1525. Internet: ehunt@u.washington.edu*

Last year, Richard Herrnstein and Charles Murray published *The Bell Curve: Intelligence and Class Structure in American Life.* Although it had more graphs than a Ross Perot speech, *The Bell Curve* made its authors' names household words, sometimes accompanied by four-letter words. Herrnstein and Murray maintained that America is splitting into the intelligent, who will move and shake society, and the less intelligent, who will be moved and shaken. They thought that the split is inevitable, because our technological society requires intelligence to run it. Finally, they said that intelligence is largely hereditary, and that numerous government programs, especially Affirmative Action, are undesirable because they amount to discrimination against the capable.

Such thoughts are not entirely politically correct. The first reactions to *The Bell Curve* were expressions of public outrage. In the second round of reaction, some commentators suggested that Herrnstein and Murray were merely bringing up facts that were well known to the scientific community, but perhaps best not discussed in public. A Papua New Guinea language has a term for this, *Mokita.* It means "truth that we all know but agree not to talk about."

The uproar over *The Bell Curve* is remarkably similar to a debate in the early 1970s. The earlier debate began when Arthur Jensen (1969) wrote that the educational enrichment programs of the Great Society were inherently limited by the immutability of intelligence and when Herrnstein (1973) claimed that differences in intelligence are largely genetic. Counterattacks followed, and by the early 1980s widely read books and articles maintained that there is no such thing as general intelligence (Gardner 1983), or that if there is it is largely a statistical artifact of the way that tests are constructed (Gould 1983), and that even if IQ exists it has little to do with life outside of a few narrow academic settings (Ceci and Liker 1986). Some of these authors have recanted (Ceci and Bruck 1994, pg. 79).

A central question in the debate is whether or not mental competence is a single ability, applicable in many settings, or whether competence is produced by specialized abilities, which a person may or may not possess independently. Almost equally important is the question of how cognitive skill, as evaluated by IQ tests, translates into everyday performance. Popular presentations on both sides of these questions leave the impression that these questions have simple answers. They do not. My goal in this essay is to discuss different theories of how intelligence is related to performance in modern society. The plural was chosen intentionally. Although we know a good deal about individual differences in human cognition, there is no monolithic, agreed-upon, all-purpose theory to organize these facts, nor is there likely to be one. There are a number of different theories that are neither right nor wrong, but are useful for different purposes.

Psychometric Views of Intelligence

In popular discussions of intelligence, including *The Bell Curve*, the term generally refers to scoring well on tests that have been developed to measure mental ability as psychologists have come to see it. I shall refer to this emphasis on test scores as the psychometric view of intelligence. Its core belief is that individual differences in human cognition can be adequately measured by performance on intelligence tests, and that intelligence itself can therefore be defined by variations in test scores, across people.

This notion was expressed most pungently when the psychologist Edwin Boring (1923), in a public debate with the columnist Walter Lippman, said that "intelligence is what the intelligence test measures." It turns out that that statement is not quite so arrogant or self-serving as it sounds. To see why we have to look at what intelligence tests are and how intelligence measures are inferred from test scores.

Although it is not always clear in our everyday use of language, scientists distinguish carefully between a conceptual variable and its operational definition—the way that it is measured. Physicists distinguish between mass as a concept and scale readings as data to be analyzed. In the best of situations there is a clearly understood link between the two. Physicists can provide a theory of the relation between a scale's movement and the mass of the object being weighed. The relation between the data for and the concept of intelligence is not at all like the relation between scale readings and mass, because in psychometrics the concept is inferred from the measuring instrument, rather than having the measurement technique dictated by the concept.

Most intelligence tests do not measure just one thing, in the sense that a scale measures only the gravitational attraction between an object and the earth. Instead, intelligence tests are made up of a number of component subtests, in which people are asked to perform different cognitive tasks. The test score is supposed to measure the common thread that runs through performance on the subtests. For instance, the widely used Wechsler Adult Intelligence Scale (WAIS) contains subtests that evaluate a person's vocabulary, short-term memory, arithmetical ability, world knowledge and several other specific skills. The Scholastic Achievement Test (SAT), which is a widely used college-screening test, and the Armed Service Vocational Aptitude Battery (ASVAB), which is used to screen military recruits, are organized in

somewhat the same way. Instead of thinking of these tests as cognitive yardsticks measuring intelligence the way a real yardstick measures length, it is better to think of an intelligence test as a sort of mental track meet, in which cognitive ability is inferred by combining subtest scores, just as athletic ability can be inferred by combining the scores in a decathlon.

This brings us to the question of how the subtest scores are to be combined. Although there is some variation from test to test, the formal basis for test combination is a statistical procedure called factor analysis. Suppose that an intelligence test consists of K subtests. (To continue the analogy to the decathlon, K is usually 10 or 12.) A person's scores on the subtests can be represented by a K-dimensional vector. The collective scores of all people in the group can be thought of as a swarm of points in a K-dimensional space. Factor analysis attempts to reduce the K-dimensional space to a smaller P-dimensional space, where $P < K$ and the axes defining the dimensions are orthogonal, or at right angles to one another. Unless the scores of two of the original tests are perfectly correlated, this always entails some loss of accuracy. The loss can be measured, so we can determine how much of the variation in the original K-space lies along a particular dimension in the reduced P-space.

To get an intuitive idea of factor analysis, imagine buying a hot dog with pimientos embedded in it. The hot dog is a three-dimensional object, so it takes three dimensions to specify the exact location of each pimiento. However, you can locate a pimiento reasonably accurately by saying where it is along the long axis of the dog. In factor-analytic terms the pimientos are the data from each person, and the three dimensions of the hot dog represent the individual tests. The long axis of the hot dog would be the first factor to be extracted and would capture most of the variation between pimiento locations. If we apply factor analysis to test
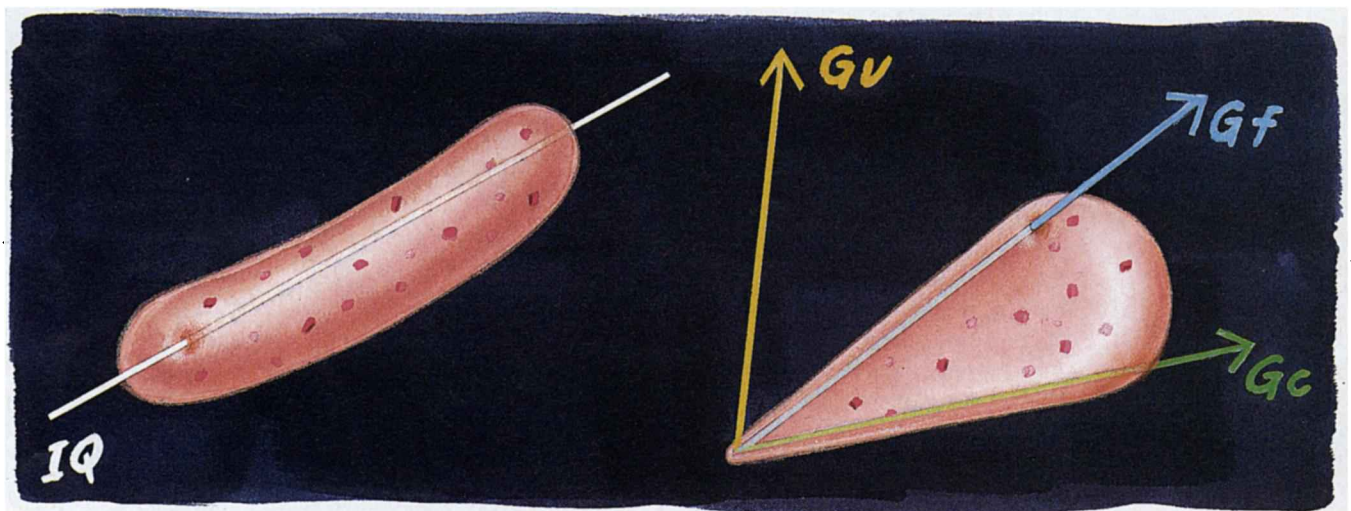
scores, instead of hot dogs, the first factor accounts for most of the variation between people just as the length of the hot dog accounts for most of the positioning of the pimientos. But instead of saying "length of hot dog," we say "general intelligence."

There are two objections to this argument. One is that when the data are reduced from the K-dimensional to the P-dimensional space, the orientation of the orthogonal dimensions in the P-dimensional space is arbitrary. To see this, consider the hot-dog example again. Although locating pimientos can be reduced from a problem in three dimensions to a problem in one dimension, the one dimension does not have to point exactly along the long axis of the hot dog. It could be rotated to any angle at all, excepting at a right angle to the long axis, and the pimientos could still be located with equal accuracy.

This fact led one critic of the idea of general intelligence, Stephen Jay Gould (1983) to argue that factor analysis is not an appropriate way of defining the variables underlying test scores, because one solution is statistically as a good as another. Gould was wrong. There are statistical methods (which were well known to specialists at the time) that make it possible to compare the goodness of fit of one factor-analytic solution to another. When these methods are applied, investigators virtually always find a highly reliable first factor. The case for general intelligence, the unitary IQ score, is far from trivial. However, there are alternative explanations for the data, based on the idea that there are different types of intelligence, even when one restricts oneself to the notion that intelligence is what the tests measure. To understand what they are, we need to delve into factor analysis a bit more.
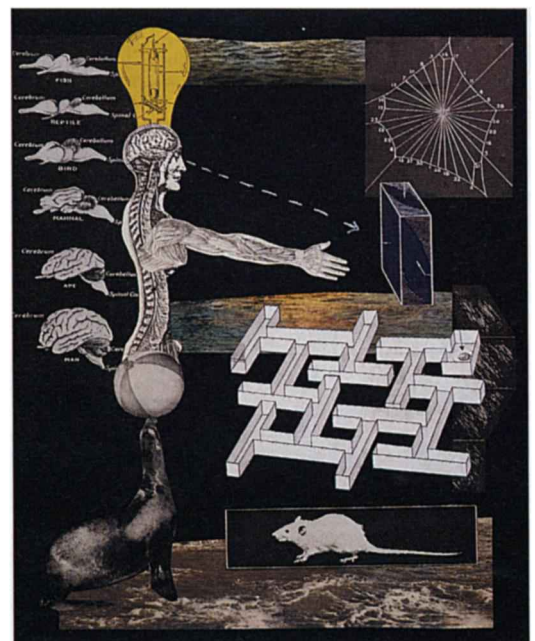
Suppose that the statistical variation in the data can be reduced from K dimensions (the original test space) to P orthogonal dimensions. This is only possible if the K original tests are positively correlated, which they virtually always are. In this case there will also be a solution in M dimensions, where $P > M > K$, in which some of the M dimensions are not orthogonal to each other. (In psychological terms, if two abilities are statistically unrelated to each other, the dimensions representing them will be orthogonal.) Now, suppose that you had some theoretical reason to believe that the data from the original K tests had been generated by two or more underlying mental factors that were statistically related to each other. Returning to the athletic example, you might want to argue that decathlon scores were determined by the strength and speed of the athletes, and that there is a statistical relationship between strength and speed. Reasoning such as this is called specifying a factor structure for the underlying abilities. Gould claimed that psychometricians could not distinguish between alternative factor structures. Today they can.

During the 1970s the Swedish psychometrician Karl Jÿeskog developed a statistical technique for evaluating the fit of a multivariate data to an arbitrary, *a priori* specified factor structure. This made it possible to compare two proposals about the structure of intelligence to data, to see which theory best fit the facts. The new methods have been applied to a number of new data sets (notably Gustafsson 1984) and have become standard in evaluating models of intelligence. In a related, highly technical but very important volume, John Carroll (1993) used somewhat different methods to reanalyze a great many important data sets that have been collected over the past 60 years. The results of these independent analyses were quite consistent. Skipping over some details, human intellectual competence appears to divide along three dimensions. Following Raymond Cattell (1971) and John Horn (1985), I shall refer to these dimensions as fluid intelligence *(Gf)*, crystallized intelligence *(Gc)*, and visual-spatial reasoning *(Gv)*. Cattell and Horn describe them as follows:

Fluid intelligence is the ability to develop techniques for solving problems that are new and unusual, from the perspective of the problem solver.

Crystallized intelligence is the ability to bring previously acquired, often culturally defined, problem-solving methods to bear on the current problem. Note that this implies both that the problem solver knows the methods and recognizes that they are relevant in the current situation.

Visual-spatial reasoning is a somewhat specialized ability to use visual images and visual relationships in problem solving—for instance, to construct in your mind a picture of the sort of mental space that I described above in discussing factor-analytic studies. Interestingly, vis-

ual-spatial reasoning appears to be an important part of understanding mathematics.

Crystallized- and fluid-intelligence measures are substantially correlated. For instance, Horn reported a study in which Gf and Gc measures were extracted from an analysis of the WAIS. The correlation between factors was 0.61. Such findings have led believers in just one intelligence to argue that Gf and Gc are simply different flavors of a general intelligence (IQ) factor. This argument cannot be answered one way or the other solely by looking at correlations between tests. However, it can be attacked by stepping outside of factor analysis and looking at how Gf and Gc measures respond to manipulations that might change mental competence. It turns out that they respond differently.

The most striking example is aging. Measures of Gf generally decrease from early adulthood onward, whereas Gc measures remain constant or even increase throughout most of the working years (Horn 1985; Horn and Noll 1994). This is not surprising. Experience counts; most of the key leadership positions in our society are held by people over 40. On the other hand, middle-aged and older people do take longer than younger people to understand new problem-solving methods and to deal with unfamiliar tasks. Age is not the only variable that can be shown to have different influences on fluid and crystallized intelligence. Alcoholism shows similar effects.

Since variables such as age, which is not itself a cognitive operation, have different influences on different types of tests, it follows that there cannot be just one ability underlying test performance. This argument moves away from the psychometric tradition, which focuses only on test scores, and towards the cognitive-psychology approach to intelligence. As the name suggests, it is derived from a more general theory about what human thought is, so a word about the general theory is in order.
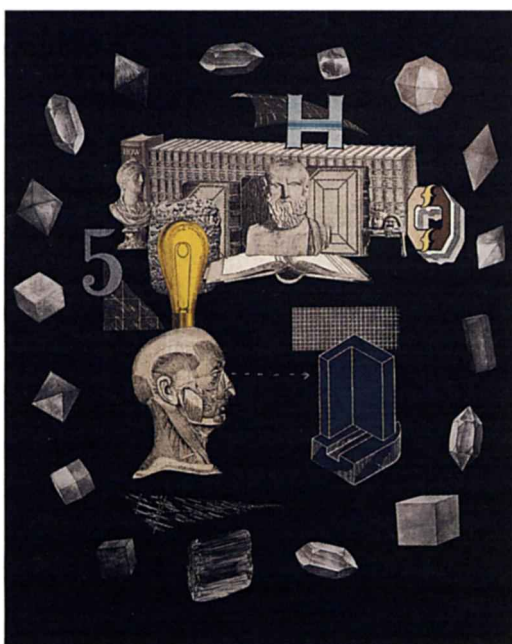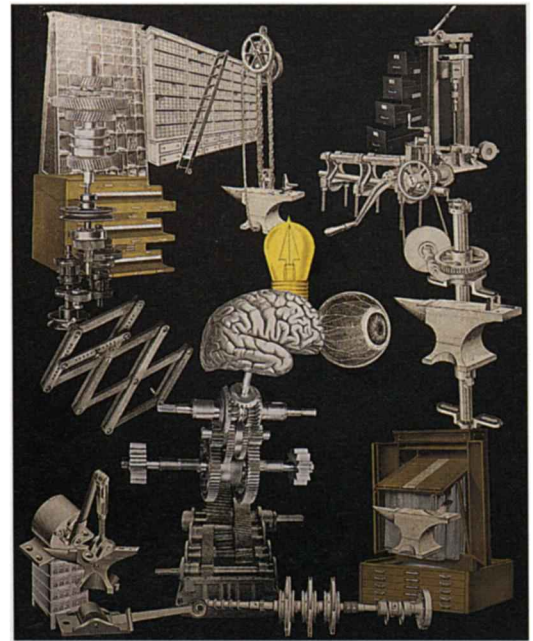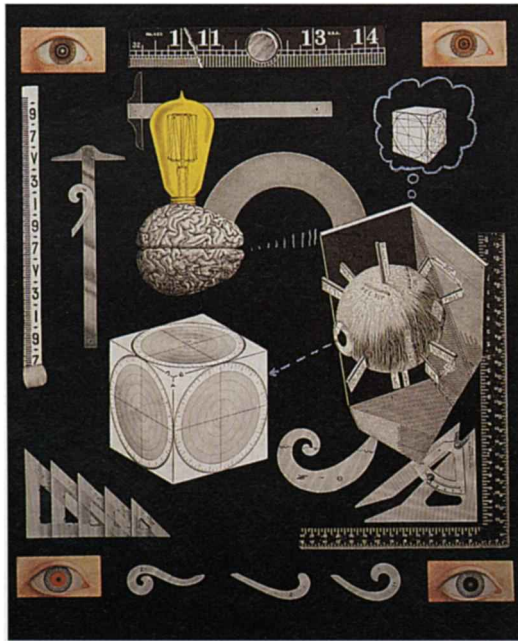
The Cognitive-Psychology View

Cognitive psychologists think of thinking as the process of creating a mental representation of the current problem, retrieving information that appears relevant and manipulating the representation in order to obtain an answer. The problem, its solution and some of the methods used to solve it are then stored for later reference. The key point in this process is creating the representation. This is assumed to require a temporary, working memory capability, which requires attention and is often a bottleneck in thought. When familiar problems are encountered the process of building an appropriate representation becomes more efficient, because previously acquired information and problem solving techniques can be used. This reduces the demand on working memory, but does not entirely eliminate it.

The cognitive-psychology view is that cognition is a process, whereas the psychometric view makes it a collection of abilities. Perhaps because it is more dynamic, the cognitive-psychology view is often seen as more appealing than the psychometric view, but it has the disadvantage of not lending itself to easy summarization. When cognitive psychologists try to characterize a person's thinking, they are not likely to use numbers to place the person in a "mental space" defined by factors derived from IQ testing. Instead they frequently use analogies to computing systems. To solve problems a computing system must have sufficient "number crunching" power to attack the problem at hand, programs that are appropriate for solving the problems the system faces, and access to the data required to solve these problems. Cognitive psychology draws an analogy between computing power, programs and data access, and the cognitive functions of being able to process ideas—any ideas—quickly and accurately, knowing how to solve certain classes of problems, and having access to the knowledge needed to solve particular problems. In psychological terms, human number-crunching is a physiological capacity, whereas knowing how to solve problems and knowing key facts are both products of learning. Each of these aspects of thought are legitimate parts of intelligence. The physiological capacities are clearly part of Gf, knowing key facts is part of Gc, and having acquired certain problem-solving strategies is a bit of both Gc and Gf. A person's capabilities are determined by the interaction between power, knowledge of how to use that power and access to required data.

The cognitive-psychology account complements the psychometric distinction between fluid and crystallized intelligence. Both accounts

The cognitive-psychology view is that cognition is a process, whereas the psychometric view makes it a collection of abilities.

In economic terms it appears that the IQ score measures something with decreasing marginal value. It is important to have enough of it, but having lots and lots does not buy you that much.

stress how a novice's performance depends on the ability to develop new problem representations (Cattell and Horn's fluid intelligence) and how with experience one shifts from problem representation to pattern recognition, by applying past solutions to present problems. Since developing a representation is more demanding of working memory and attention than pattern recognition is, learning to do an intellectual task will generally be harder than doing it. The theory also implies that people who do well on tests of fluid intelligence should have a large working-memory capacity, and indeed, they do (Carpenter, Just and Shell 1990).

When cognition is viewed this way it is not surprising that IQ tests, and especially fluid-intelligence tests, are associated with academic performance. By definition students are novices. So are apprentices in workplace settings. Data from the military (Wigdor and Green 1991) have shown that performance on the Armed Forces Qualification Test (AFQT), which is used to screen military recruits, has a strong relation with performance on the job in the first few months. After two years the relation is reduced, but not negligible. Similarly, the Department of Labor's General Aptitude Test Battery (GATB) has been shown to be less valid for older than for younger workers. This is consistent with laboratory studies and theoretical analyses in cognitive psychology, all of which show that the experience reduces but does not eliminate the relation between general intelligence and performance (Ackerman 1987).

Nonlinearities in Intelligence

Most of our everyday measurements are linear measurements. A linear measurement is one in which a constant interval means the same thing at any point on the scale. For instance, adding one inch to a six-foot board produces the same change in length that adding one inch to a five-foot board does. We are so familiar with linear measurements that we often assume that the properties of linear measurements apply to any characteristic that is described by numbers. That is not so, and the erroneous assumption can be particularly confusing when we deal with intelligence.

In psychometric theories intelligence is calculated by determining a person's standard score on an IQ test. The standard score is the deviation of a person's absolute score of a test from the mean test score of a reference population, divided by the standard deviation (a measurement of the variability of scores in the reference population):

$$z_i = \frac{(x_i - \mu)}{\sigma}$$

where x_i is the ith person's score in absolute units (usually the number of correct answers on a test) and μ and σ are, respectively, the population mean and standard deviation. If this equation were applied strictly, a person of exactly average intelligence would have a score of zero, and people with below-average intelligence would have negative scores. Since the ideas of zero and negative intelligence do not seem reasonable, it is conventional to report IQ scores by rescaling standard scores, using the equation

$$IQ = 15z + 100$$

This gives the person of average intelligence a score of 100. This equation is simply a scaling convention; the real definition is contained in the first equation, which makes the standard deviation the unit of scoring. Herrnstein and Murray refer to the standard deviation as "like an inch," but it is not. The standard deviation is determined not by the absolute values of the

scores in a population, but rather by the extent to which one score is likely to be different from another. In addition, the zero point of the IQ scale (IQ = 100) is determined by the population mean, not by a definition of "average intelligence" in terms of intellectual performance. Therefore the IQ score of an individual is a relative score, compared to the mean and variability in the reference population, rather than an absolute measure of mental competence. If we measured height the way that we measured IQ, a six-foot, six-inch man would have a standard score of somewhat greater than 2, in the North American male population. The same person would have a standard score of about 0 if the reference population were professional basketball players.

The distinction between the relative and absolute definitions of intelligence becomes important when we consider the relation between IQ, defined by standard scores, and various dependent measures, such as school achievement and workplace performance. Suppose a psychometrician records the job performance and intelligence-test scores of a group of workers. The relationship would be expressed by this equation, where B is the regression coefficient, or the rate at which job performance changes as IQ changes:

$$\text{job performance} = \text{average job performance} + B \cdot \text{IQ}$$

B is calculated to make predictions as accurate as they can be. The actual degree of accuracy is measured by the correlation coefficient , which varies from 0 (no accuracy at all) to 1 (perfect prediction). Determining the regression and correlation coefficients from a given set of data is straightforward. The problem comes when an extrapolation is made to new situations, where some data points lie outside the range of IQ units observed in the original study. An example might be extrapolating the grade-IQ relationship observed in high-school students to grade-IQ relations among college students. Such extrapolations implicitly assume that IQ scores are linear measures of the intellec-

tual traits that they are supposed to measure. This is not true. Suppose that a person in his 20s suffered a brain injury or infection that reduced his IQ score by 20 points. (Such things are possible.) If he were a medical or law school student with an original IQ of 140, he would probably still complete his coursework, though perhaps with not quite so high a class rank as before. If the person were a blue-collar worker with an original IQ of 80 he would, at IQ 60, have a substantial risk of homelessness, poverty and a number of other serious social problems.

The issue of nonlinearity applies to the very definition of intelligence, and in particular to the question of whether there is one type of intelligence or several. Suppose that general intelligence is equally important at all levels of mental competence. In this case the results of a factor-analytic study of test scores, based on data from people with high levels of intelligence, should be similar to the results of a study based on data from people of lower absolute levels of intelligence. Historically there have been suggestions that this is not so. The general-intelligence model was first developed by Charles Spearman (1904, 1927), based on analysis of test results from English schoolchildren. In 1938 L. L. Thurstone challenged Spearman's conclusion because he found very little evidence for general intelligence in a sample of University of Chicago undergraduates. It was observed at the time that the discrepancy might have arisen because Spearman and Thurstone had taken data from people of widely different intellectual levels, which would be evidence that intelligence changes qualitatively as the level of mental competence changes. However, the results were not definitive because Spearman and Thurstone had used different tests.

An important study by Douglas Detterman and Mark Daniel (1989) showed that the relations between subtests do change as the level of scores changes. Among other things, Detterman and Daniel examined correlations between subtests of the WAIS and found higher correlations between subtest scores for people with below-average IQ than for people with

a linear scale – if A=B, then d=b

a non-linear scale – A=B, but a≠b
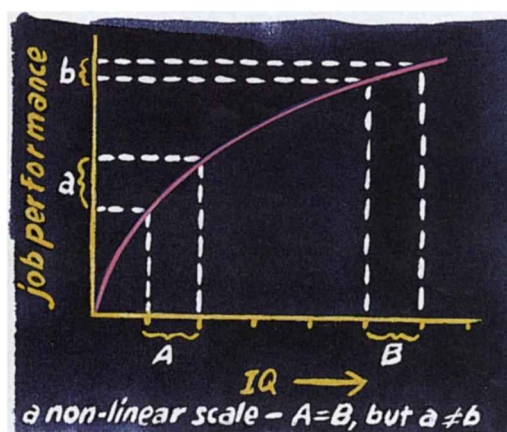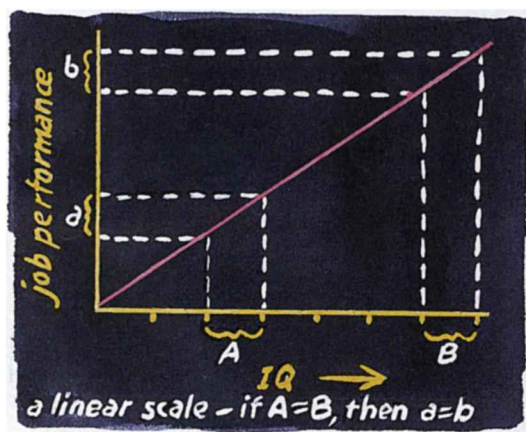
above-average IQ. David Waller and Derek Chung and I found the same thing when we analyzed the ASVAB scores that Herrnstein and Murray used in *The Bell Curve* to determine the relation between IQ and various indicators of social adjustment. It appears that general intelligence may not be an accurate statement, but general lack of intelligence is!

The conclusion that the relation between different indices of mental competence depends on the general level of competence is not consistent with psychometric approaches, but it is consistent with the cognitive-psychology approach. Recall that the cognitive-psychology approach assumes that mental competence is produced by a cascade of progressively more refined abilities, moving from information processing to problem-solving techniques to knowledge possession. It follows that problems at the information-processing level will be general, whereas potentials established at higher levels will be specific. In fact, Detterman and Daniel did find that the relation between information-processing measures and intelligence-test performance is higher at low levels of intelligence. Similar observations have been made by scientists who have studied very high-level performance, in fields ranging from physics to literature. A certain amount of intelligence seems to be needed to gain entry to an intellectually demanding field, but beyond that point success is determined by the effort put into the job, social support, and just sheer experience. (See Ericsson, Krampe and Tesch-Romer (1993) on expertise, Simonton (1984) on creativity, and Gardner (1993) for some interesting biographical data.)

In economic terms it appears that the IQ score measures something with decreasing marginal value. It is important to have enough of it, but having lots and lots does not buy you that much. My regrets to Mensa, but that is the way things are. Nonlinearity becomes important when we ask a key question raised by Herrnstein and Murray: What is the relation between intelligence and workplace performance?

How Important Is Intelligence?

No one would worry about who has intelligence, or why, if it did not matter. Indeed, one of the claims made by the opponents of testing in the 1960s and 1970s was that intelligence tests just measured academic performance, and that even there they did not do a good job. One of Herrnstein and Murray's major contributions has been to expose this bit of Mokita. Intelligence, as measured by the tests, really does matter in both school and workplace, although it may matter in somewhat different ways than *The Bell Curve* suggests.

To argue that IQ is a determinant of economic outcomes, Herrnstein and Murray re-

lied on two sources of evidence. One was the recent literature, and especially John Hunter's (1986) summary of the relation between IQ scores and workplace performance. The other was their own analysis of data from the National Longitudinal Survey of the Labor Market Experience of Youth (NLSY). The NLSY is a Department of Labor survey that has followed over 12,000 participants since 1979. The respondents are now in their late 20s and early 30s. Early in the survey many participants took the Department of Defense's ASVAB test. Herrnstein and Murray used the AFQT score, which is derived from the ASVAB subtest scores, as a measure of IQ. They then related IQ to subsequent life events, such as being employed or being below the official poverty line.

Hunter reviewed studies of the relationship between job performance and scores on the General Aptitude Test Battery (GATB), a Department of Labor test which was widely used until the late 1980s, when the testing program became embroiled in a controversy over its fairness to minorities. The GATB was withdrawn as a political rather than a scientific decision. After a detailed statistical analysis, Hunter concluded that the "true" relation between intelligence and job performance in the population is about 0.5. This conclusion depended heavily upon extrapolating relationships beyond the data, which assumes linearity. A National Science Committee reviewing the GATB argued that Hunter should have used the observed correlations, which were almost all in the 0.2 to 0.3 range. The truth probably lies between these estimates, providing that the extrapolation is to comparable jobs (Hunt 1995). And that is an important qualification.

The GATB was designed to screen applicants for entry-level jobs in blue-collar and lower-level white-collar occupations. In terms of averages (something that is well established), we are talking about occupations where the mean IQ is in the 90–110 range, which covers about half of the population. But recall that as intelligence goes up cognitive abilities become more differentiated. Also, as experience goes up the IQ-performance connection gets weaker. These factors would lead to a reduction in IQ-performance relations within higher-level job classifications, and when dealing with experienced and older individuals. (In fact, the GATB is known to be less accurate in predicting the performance of older workers.)

The qualification within a job class is also important. There are quite high correlations between the socioeconomic status of a job and the mean IQ of the jobholders. Truck drivers average slightly under 100, while high-paid professionals, such as doctors and lawyers, have averages of 125 or above. It is sometimes asserted that this is because general intelligence is need-

ed to obtain the educational certification required to qualify for a job, but is less important to on-the-job performance. There is evidence for this. Military and civilian studies have found that IQ tests are better predictors of performance when people are in training programs than when they are on the job itself. After people are on the job, correlations are higher between IQ and tests of job knowledge than between IQ and on-the-job observations of performance. However, none of the correlations vanish.

IQ does not predict all aspects of job performance. In an extensive study of enlisted personnel (Campbell, McHenry and Wise 1990), the Army found that it was useful to distinguish between what might be called ability aspects of performance, which includes such things as knowledge of one's job requirements and the ability to operate machinery required in the job, and motivational aspects, which include cooperating with colleagues, showing initiative and leadership. The ASVAB did a good job of predicting the ability aspects but had almost no relation to the motivational aspects. This is not surprising, but it does make any focus on a unitary index of job competence seem simplistic.

In summary, it appears that IQ is an important factor in getting into a job or profession, but is less important (although not negligible) once you have learned to do the job. Further improvement is then achieved by acquiring experience, rather than improving upon an abstract knowledge of what the job requires.

Untangling Social Variables

If we can predict good things, however imperfectly, for someone with a high IQ score, what can we predict for a person with a low score? People with criminal records, people who are below the official poverty line, and people who are receiving aid for dependent children tend to have low IQ scores. Based on their analysis of the NLSY data base, Herrnstein and Murray argued that IQ causes these problems, because AFQT scores are often the best single predictor of a person's social troubles.

People who are below the poverty line are likely to simultaneously have low IQs (on the average) and poorer than average health, and to come from parental families of low socioeconomic status (SES). What is causing what? The question is hard to answer, partly because of the difficulty of the statistical analysis and partly because most social problems have multiple causes. Young adults on welfare may be there because of a combination of low intelligence, lack of education and limited familial support.

In preparing their book, Herrnstein and Murray used a technique called logistical regression to attack the statistical problems. They first defined a binary social variable, such as having an income under the official definition of poverty, and then looked at the relation between the probability that a person will be on the bad side of this variable as a combined function of various predictor scores, such as IQ (defined by the AFQT), SES, and education. Because of mathematical problems, it is not possible to look at the probability of, say, poverty status, directly. Instead they calculated a regression equation. In this equation p is the probability of being in poverty. A logarithmic expression based on p is related to IQ, SES, education (ED) and so forth by the regression coefficient for each (the B terms).

$$\ln\left(p/(1-p)\right) = A + B_{IQ}(IQ) + B_{SES}(SES) + B_{ED}(ED) + \ldots$$

If all variables are expressed as standard-score units, you can determine the relative importance of each variable as a predictor by comparing the regression coefficients. For instance, in the case of poverty status the regression coefficient for IQ is -0.84 and the regression coefficient for SES is -0.33. This tells us that the risk of poverty goes up as IQ and parental SES go down, and that, since the absolute value of the IQ regression coefficient is greater than the absolute value of the SES regression coefficient, the risk of poverty is more sensitive to changes in personal IQ than to changes in parental SES.

Results like this are ubiquitous in the NLSY data. IQ is the best predictor of being below the official poverty line, dropping out of high school and receiving aid for dependent children. IQ and SES are about equal in predicting risks of long-term unemployment and of divorce. Since the publication of *The Bell Curve*, and possibly inspired by it, there have been a number of privately circulated alternative analyses of the NLSY data. All the ones that I have seen show that, although you might change the exact numbers reported by Herrnstein and Murray a bit, intelligence is a substantial predictor of indicators of social problems.

But just how substantial, and how should a prediction based on intelligence be related to a prediction based on other factors? This is a hard question to answer, because of the complicating factors of nonlinearity and collinearity. Recall that nonlinearity means that a relation is not the same at all levels of the predictor (IQ). Understanding nonlinearity is always difficult. The problem is compounded because, in this case, the regression coefficients are not for the risk of a social problem; they are for the logistic function of that risk. This function is not intuitive to most people. Collinearity refers to the fact that the predictor variables—IQ, SES, education and a number of other possible predictors—are themselves highly correlated. In the NLSY data, for instance, the correlation between IQ and SES is 0.55, which is about as high as the correlation between adult height and weight.
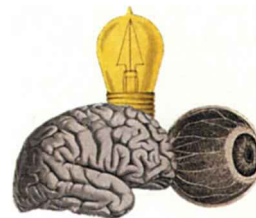
Once a person's scores are in the moderate SES or moderate cognitive ability ranges the relation between poverty, IQ and parental socioeconomic status virtually vanishes.

Within the framework of the psychometric defintion, the evidence is quite clear that intelligence is substantially inherited.

The graph below shows how these effects combine in the NLSY. This figure is a three-dimensional view of the relation between the probability of being in poverty status, represented by color; IQ (the horizontal axis); and SES (the vertical axis). The figure shows both the nonlinearities and the collinearity of these data. For anyone of above-average intelligence or high parental SES, the probability of being in poverty status is very low indeed. This is indicated by the large black area in the figure. Furthermore, in this distribution people with moderate or better SES and very low intelligence, or moderate to better intelligence and low SES, are not likely to exist. (Note that the figure is not a square.) The red "hot spot" might be thought of as a danger zone in which relatively high probabilities of poverty status are associated with the combination of the bottom 15 percent of the intelligence and parental SES distributions. This does suggest a troubling, cyclical relation between these variables. But once a person's scores are in the moderate SES or moderate cognitive ability ranges the relation between poverty, IQ and parental SES virtually vanishes.

Waller, Chung and I have developed a number of similar analyses for other "at risk" variables in the NLSY data set, such as health problems and prolonged unemployment. No single picture emerges. What is clear, though, is the need to consider nonlinearity and collinearity in each case. Even after this is done, intelligence test scores in the bottom 15 percent (roughly an IQ of 85 or below) almost always indicate that a person has a substantial risk of encountering problems in our society. It is important to remember that this is a statistical statement, whereas at the individual level nonstatistical interactions are involved. There are undoubtedly many cases in which a person with low parental SES inherits genetic limitations in IQ, and IQ score is indicative, on average, of the extent to which a person can

benefit from education. There are other cases in which limited family support or limited educational opportunity may restrict a person's intellectual potential, even when a person is highly motivated to succeed. Statistics cannot tell us to what extent any of these variables is operating in an individual case. All statistics can tell us is how many such cases to expect in the population.
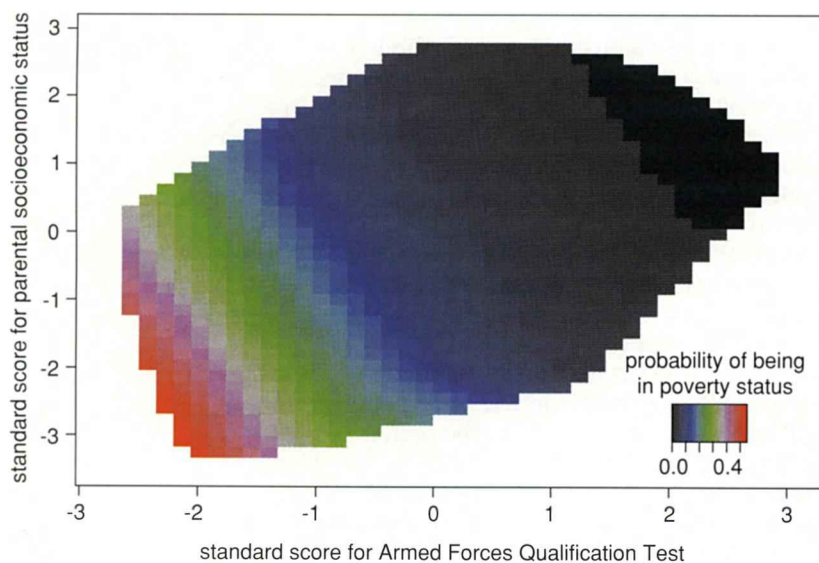
We once again see that the data are more easily explained by the cognitive-psychology view of intelligence as an interacting process than from the psychometric emphasis on linear relationships. From a cognitive-psychology perspective, low IQ might cause social problems, because of the failure of some general component of cognition, but once beyond a given level of ability people would be able to cope with the general society adequately. (Anthropologists will hardly be surprised to find that most people are able to operate in their own cultures!) Social problems could arise, though, if the threshold for doing well in society were set so high that a substantial number of people could not meet it. This topic will appear again when we look at the interaction between scientific facts and public policies.

Can Cognitive Abilities Be Improved?

Because expressed intelligence must be drawn out from innate ability, through cultural experiences, it is natural to ask whether certain cultural experiences, including education, can improve intelligence. Some social programs have had this as an explicit goal. It is also natural to ask whether societies can improve intelligence by altering the physical environment—for instance, through programs to improve nutrition or the family environment. Finally, whether or not intelligence, as measured by tests, is subject to improvement, there remains the question of whether cognitive *competence* can be manipulated.

These questions have been looked at in three ways: in statistical and historical comparisons of cultures, from within our own culture's experience and from the viewpoint of statistical and theoretical biology. They are at the core of the debate reignited by Herrnstein and Murray, who argue that competence in today's workplace is determined by IQ, that IQ is determined by inheritance and that since IQ is resistant to change, social programs that rely on changing or disregarding IQ are misguided and even counterproductive.

If we take a cross-cultural perspective, there is evidence that broad characteristics of a society can influence reasoning, probably by placing a value on the practice of certain intellectual skills. Literacy is associated with an appreciation for abstract reasoning, which is of considerable importance in a technologically oriented society. Nonliterate, traditional cul-

standard score for parental socioeconomic status

probability of being in poverty status

0.0   0.4

standard score for Armed Forces Qualification Test

tures seem to place more weight on reasoning based on memory and personal experience. These observations, though, are of limited importance for the study of variations in intelligence within our own society, where minimal literacy is virtually universal.

There is some indication that intelligence levels have changed over time within Western cultures. Flynn (1987) observed that the absolute scores on widely used tests of abstract reasoning *(Gf)* have increased in North America and Europe since World War II. Interestingly, scores on tests that are designed to evaluate cultural knowledge and problem-solving techniques *(e.g.,* the SAT) declined over the same period. Although the reasons for these changes are not known, the fact that they have moved in the opposite direction is further evidence for distinguishing between intelligence as an abstract problem-solving ability from intelligence as an ability to attack culturally relevant problems.

When we move from comparisons across cultures and across time to our own society, we find surprisingly little evidence for influences of cultural experiences on intelligence—once again, as measured by intelligence-test scores—in spite of many efforts to find such effects. Two well-documented findings capture the gist of the results. Studies of adopted children have repeatedly shown that the IQ of the biological parent is a better predictor of the child's IQ than is the IQ of the adopting parent, even when adoption is virtually at birth. Consistent with this observation, the quality of home or school environments appears to have relatively little relation to permanent changes in test scores, once one has taken account of the correlation between genetic and social variables. Put a slightly different way, genetic predictions based on parental or sibling IQ can account for IQ variability in children, after social factors have been taken account of, but social factors are not related to children's IQ after genetic variability has been accounted for (Scarr, in press).

Within the framework of the psychometric definition, in fact, the evidence is quite clear that intelligence is substantially inherited. Behavior-genetics studies have shown repeatedly that IQ scores behave as if between 40 and 80 percent of the variation in intelligence, across individuals, can be accounted for by genetic variation. The exact value does not matter. Identical (monozygotic) twins who are adopted at birth and raised apart will resemble each other in IQ more than fraternal (dizygotic) twins raised together. Genetic heritability of IQ is a major determinant of whatever is behind the IQ scores.

Genetic heritability has become entangled with racial and ethnic issues each time the national intelligence debate has flared up. Gaps in intelligence-test scores among groups exist;

Herrnstein and Murray, like Jensen before them, posit a genetic explanation. Many social activists have responded by denying the tests' validity in minority groups. The facts in this debate are pretty clear, but the explanation for the facts is not.

Numerous studies have found that in the United States the average IQ score in samples of blacks and Latinos is about one standard-deviation unit below the average score for whites and Asians. This means that the median black score is exceeded by 87 percent of whites. There is, at best, marginal evidence showing that the tests do not predict minority academic performance as well as they predict majority performance. With a few exceptions (primarily involving language tests in Latinos) test items that appear to have the least cultural bias show some of the largest ethnic-group differences. Herrnstein and Murray asserted that the tests are equally valid for minorities and majorities; although too strong, this statement is closer to the truth than the claim that the tests are totally invalid. This does *not* mean that the differences in IQ scores between ethnic groups are genetic in origin. In our society ethnic status and social variables that might correlate with intelligence are highly confounded. Therefore the currently available data do not discriminate between genetic and nongenetic explanations. We do not know whether ethnic-group differences are innate or not. Given the complexities of the situation, not the least of which is defining what ethnic group a person belongs to, we should perhaps let the issue go at that.

IQ and Cognitive Skills

The view is different as soon as one steps outside psychometrics. The sociologist Christopher Jencks (1992) has observed that genetic explanations that stop with a heritability coefficient are unsatisfactory because they do not specify how intelligent behavior is produced. No one inherits an intelligence-test score in the sense that one inherits eye color. What must be inherited is a physiological capacity for paying attention, learning and reasoning that allows us to extract from our experiences the knowledge and problem-solving techniques required to solve test problems. We have very little idea about what these physiological mechanisms might be, especially insofar as they are related to variation in abilities within the normal range of intelligence. (There is a considerable knowledge of physiological problems associated with specific types of mental retardation.)

Whichever model they adopt, psychologists have been frustrated in the search for ways to enhance cognitive function. Research has shown how we might lower a person's intelligence by physical intervention, but not how to

The economic issue is what skills people possess, not what their IQ scores are.

improve it. There are drugs that produce brief improvements in specific cognitive functions, such as memory or attention, but the intelligence pill is nowhere in sight. And although nutrition might be thought to be a significant effect, there is at best marginal evidence for nutritional effects within the range of nutrition encountered in the developed world.

Even if we do not know how to improve intelligence, as indicated by the test scores, the economic issue is what skills people possess, not what their IQ scores are. We may not be able to destroy the linkage between IQ scores and the relative possession of cognitive skills (and it is not clear why we would want to), but improved education and training can raise the average achievement of all students.

A study by one of my colleagues (Levidow 1994) showed this in a controlled way. High-school students were given a test of fluid intelligence. They then took a year-long problem-solving–oriented course in elementary physics. The IQ test did indeed predict how much physics the students learned. At the end of the year they took an equivalent IQ test. Their IQ scores had not changed a whit. Furthermore, the IQ test did predict the relative standings of the students on the final examination. However, all students had learned a great deal of physics, as evidenced by comparisons to national standards. IQ may not have been changed, but cognitive competence, in the sense of the problems the student could solve, was increased.

Levidow's study involved a carefully monitored educational program. Could similar increases in skill be obtained just by putting more effort into education? In 1994 the New York City school system, at the insistence of their new chancellor, required that virtually all 10th-grade students take science courses that previously had been taken by only half the students, usually the more able ones. Enrollment jumped from 20,000 to 48,000 students. Failure rates went up, from 13 percent to 25 percent. Pessimists can point to this as a consequence of trying to teach hard topics to less-intelligent students. There is probably some truth to this. But more than twice as many students successfully completed science courses in 1994 than in 1993.

I have just cited examples of programs that achieved success by one measure, which happens not to be IQ scores. Herrnstein and Murray cited different examples to buttress their conclusion that programs intended to enrich children's intellectual experiences, such as Head Start, have failed. This has serious policy implications, because enrichment programs are generally targeted toward children who, as a statistical group, have low IQ and are considered at risk for school failure. Saying that the programs have failed is a bit strong, because

the programs certainly should not be judged solely by their effect on children's IQ scores, and perhaps not even solely upon children's school records. But by these measures it is clear that enrichment programs have not been nearly as successful as it was hoped that they would be when they were initiated in the 1960s and early 1970s.

What measures are appropriate to judging such programs? In our society the labor market supplies the yardstick. Herrnstein and Murray maintained that changes in our society are increasing the value of intellectually demanding occupations, relative to the value placed on less intellectually demanding ones. For example, they would argue that in modern times the values to society of computer-system designers and bank-portfolio managers have increased relative to the values of bookkeepers and tellers. They are not the only ones to have made this observation. Secretary of Labor Robert Reich (1991) has described the ascendancy of the "symbol analyst," the person whose expertise is in dealing with abstract models of the world rather than dealing with it directly. The evidence for this trend is overwhelming, and all indications are that it will be accelerated by technological changes that are clearly on the horizon (Hunt 1995).

The trend has implications for economic investment in education. During the 1960s and 1970s, and to a considerable extent today, special funds were made available to deal with the "at risk" student, where there was a greater expectation of educational failure. Much less was spent on funding for gifted students. Herrnstein and Murray argue that this is a poor investment policy, on the grounds that education produces a greater added value for society when applied to the top student than when applied to the bottom one. They also argue that because IQ is the driving force in workplace success, and because little can be done to change it, little can be done to change the situation at the bottom.

Given the evidence for increasing economic value for highly educated, skilled workers, this is not unreasonable. A good case can be made for investing more in the development of high-level skill than we do now. The United States charges tuition to university students who, in other industrial countries, would receive stipends as part of an effort to improve national human resources. Two qualifications have to be added. One is that because of the nonlinearities between intelligence and performance, as documented above, it is not clear that the gains from the cultivation of high-level skills would be as great as *The Bell Curve* suggests. The other is that because SES is positively correlated with intelligence emphasizing the development of upper-level intellectual skills does tend to make the fortunate more fortunate. The economic ad-

Investments that improve the efficiency of training and education will have larger and larger payoffs as the technological sophistication of society increases.

vantages of the investment have to be weighed against our society's general disinclination to support the privileged.

When it comes to programs to improve cognition generally, there is little room for argument. We need to increase competence at all levels because the increasing technological nature of our society has both increased the opportunities available to the capable and increased the penalties for not being able to keep up. Consumer credit is a good example; new banking technologies have provided the average citizen with an opportunity for leveraged investment that were previously open only to the wealthy. (This is what a credit card is!) Managing the opportunity requires a good bit of sophistication, so consumer debt is a problem. The cognitive skills needed to be a fully functional member of our society are clearly on the rise. Once again, intelligence is more closely linked to acquiring these skills than to exercising them once they are acquired. Therefore investments that improve the efficiency of training and education will have larger and larger payoffs as the technological sophistication required to function in society increases.

Intellectual Resources in the Workforce

Facts about intelligence are relevant to policy in another area: the question of how society should use those resources that it already has. Affirmative-action programs are now on the political chopping block, and the question raised by Herrnstein and Murray—Do they discriminate against the capable, and thereby squander the nation's intellectual resources?— is squarely in front of us.

From a narrow perspective, if the payoff for performance is highest at the top end of intellectual demands, we should be zealous about ensuring that the most demanding, generally best paid, jobs do in fact go to the most competent. To the extent that IQ scores indicate who these people are, we should pay a premium for intelligence. This policy, which Herrnstein and Murray (and others) advocate, has an unfortunate side effect. At the present time assignment of jobs solely on the basis of performance predictors, such as skills tests, would result in marked underrepresentation of minorities in high-level job classes. This, in itself, would create a costly division in society, because the ethnic groups involved would understandably refuse to accept this outcome as just.

The only way out of this situation is to make major investments in training and education in the affected communities, so that the distribution of workforce skills becomes more equitable across ethnic groups. There is also a good deal of evidence that successful investment must include participation and support by the minority communities themselves. Simply admitting more minority-group members to present programs does not work. In fact, there is evidence that some such efforts have amounted to certification that minority group members have passed through an educational program without a concomitant emphasis on performance. A recent survey of workplace skills showed that blacks with graduate-school experience have, on the average, writing and computational skills equivalent to whites who have only a community-college education (Kirsch et al. 1993). The issue is the changing of skill levels, not certification levels!

The Bell Curve leaves the impression that nothing can be done because of immutable IQ differences. This position goes beyond the evidence. In fact, Herrnstein and Murray admit that some educational improvement programs that they regard as far too expensive to be feasible nationwide have been effective. The decision about whether a program is "too expensive" or not is a matter of political rather than scientific judgment.

As this essay has shown, our knowledge of intelligence has been extracted from complex statistical relationships. Queen Victoria's Prime Minister, Benjamin Disraeli, said, "There are lies, damned lies, and statistics." What social policies are dictated by selected facts about intelligence depends on who is doing the selecting. Besides, while social policies are certainly constrained by scientific findings, it is seldom the case that findings in the social sciences will dictate just one policy.

Variations in intelligence have always been with us. How important they are depends on the technological level and social organization of society. The "village idiot" was a stock figure in medieval and early industrial stories. In pre-industrial days, though, an able-bodied person, living in a tightly knit society where economic, extended family and social roles merged, may have been able to be a contributing member of society. In fact, in such societies most of the brighter members of society may not have been able to divorce themselves from the problems of dealing with such individuals, so that it was to their advantage to see that everyone could cope. This probably became less true as agrarian societies were replaced by industrial ones. Today we live in a society where economic roles dominate other roles, where the extended family is reduced to an exchange of Christmas cards with cousins (and even ex-spouses) and where the movers and shakers of society can, indeed, afford to remove themselves from the moved and shaken. There are fascinating questions here for those interested in the intersections between sociology, economics, anthropology and cognitive psychology. We do not have the answers yet. We may need them soon, for policy makers who rely on Mokita are flying blind.

Today we live in a society where economic roles dominate other roles.

Acknowledgment

Bibliography

Ackerman, P. 1987. Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin* 102(1):3–27.

Campbell, J. P. 1990. An overview of the Army Selection and Classification Project (Project A). *Personnel Psychology* 43:231–239.

Campbell, J. P., J. J. McHenry and L. L. Wise. 1990. Modeling job performance in a population of jobs. *Personnel Psychology* 43:313–333.

Carpenter, P. A., M. A. Just and P. Shell. 1990. What one intelligence test measures. A theoretical account of processing in the Raven Progressive Matrix Test. *Psychological Review* 97 (3):4-4-431.

Carroll, J. B. 1993. *Human Cognitive Abilities.* Cambridge: Cambridge University Press.

Cattell, R. B. 1971. *Abilities: Their Structure, Growth, and Action.* Boston: Houghton Mifflin.

Ceci, S. J., and M. Bruck. 1994. The bio-ecological theory of intelligence: A developmental-contextual perspective. In *Current Topics in Human Intelligence,* ed. D. K. Detterman. Volume 4: Theories of Intelligence. Norwood, N.J.: Ablex.

Ceci, S. J., and J. Liker. 1986. Academic and nonacademic intelligence: An experimental separation. In *Practical Intelligence: Nature and Origins of Competence in the Everyday World,* ed. R. J. Sternberg and R. K. Wagner. Cambridge: Cambridge University Press.

Detterman, D. K., and M. H. Daniel. 1989. Correlations of mental tests with each other and with cognitive variables are highest in low IQ groups. *Intelligence* 13:349–360.

Ericsson, K. A., R. Th. Krampe and C. Tesch-Romer. 1993. The role of deliberate practice in the acquisition of expert performance. *Psychological Review* 100(3):363–406.

Gardner, H. 1983. *Frames of Mind: The Theory of Multiple Intelligences.* New York: Basic Books.

Gardner, H. 1993. *Creating Minds.* New York: Basic Books.

Gould, S. J. 1983. *The Mismeasure of Man.* New York: Basic Books.

Gustafsson, J. E. 1984. A unifying model for the structure of intellectual abilities. *Intelligence* 3:179–203

Herrnstein, R. J. 1973. *I.Q. in the Meritocracy.* Boston: Little, Brown.

Herrnstein, R. J., and C. Murray. 1994. *The Bell Curve: Intelligence and Class Structure in American Life.* New York: The Free Press.

Horn, J. L. 1985. Remodeling old models of intelligence. In *Handbook of Intelligence. Theories, Measurements, and Applications,* ed. B. B. Wolman. New York: Wiley. Pg. 267–300.

Horn, J. L., and J. Noll. 1994. A system for understanding cognitive capabilities: A thoery and the evidence on which it is based. In *Current Topics in Human Intelligence,* ed. D. K. Detterman. Volume 4: Theories of Intelligence. Norwood, N.J.: Ablex.

Hunt, E. 1995. *Will We Be Smart Enough? A Cognitive Analysis of the Coming Workforce.* New York: Russell Sage Foundation.

Hunter, J. E. 1986. Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior* 29:340–362.

Jencks, C. 1992. *Rethinking Social Policy: Race, Poverty, and the Underclass.* Cambridge, Mass.: Harvard University Press.

Jensen, A. R. 1969. How much can we boost IQ and scholastic achievement? *Harvard Educational Review* 39(1):1–123.

Kirsch, I. S., J. Jungeblut, L. Jenkins and A. Kolstad. 1993. *Adult Literacy in America.* Washington: National Center for Educational Statistics.

Levidow, B. B. 1994. The effect of high school physics instruction on measures of general knowledge and reasoning ability. Unpublished Ph.D. Dissertation, U. of Washington.

Reich, R. 1991. *The Work of Nations: Preparing Ourselves for 21st Century Capitalism.* New York: Knopf.

Scarr, S. In press. Behavior genetic and socialization theories of intelligence: Truce and reconciliation. In *Intelligence, Heredity, and Environment,* ed. R. J. Sternberg and E. G Rigorenko. Cambridge: Cambridge University Press.

Simonton, D. K. 1984. *Genius, Creativity, and Leadership: Historiometric Inquiries.* Cambridge, Mass.: Harvard University Press.

Spearman, C. 1904. General intelligence, objectively determined and measured. *American Journal of Psychology* 15:201–293.

Spearman, C. 1927. *The Abilities of Man.* London: MacMillan.

Thurstone, L. L. 1938. *Primary Mental Abilities.* Chicago: University of Chicago Press.

Wigdor, A. K., and B. F. Green, Jr. 1991. *Performance Assessment in the Workplace.* Washington: National Academy Press.