Chapter 6
The *g* Factor and the Design of Education

Arthur R. Jensen
University of California, Berkeley

A scientific theory of intelligence, in the true sense of the term, cannot be translated directly into a prescription for practice. In any case, a real theory of intelligence does not yet exist. If it did exist, it would describe the anatomical, neural, and chemical processes in the brain that govern the types of behavior that are claimed to characterize intelligence. The major aim of present-day cognitive neuroscience is to achieve this kind of understanding of how brain mechanisms produce intelligent behavior, but as yet we have no comprehensive account of how the brain does this. After all, the human brain is the most complex structure known to modern science, and the technology needed for probing the living brain in action, such as the measurement of evoked electrical potentials by electroencephalography, or of the metabolic rates in specific regions and structures of the brain by positron emission tomography (PET) scan, or mapping localized functions by magnetic resonance imaging (MRI), are all of fairly recent origin in the history of scientific instrumentation. Hence, scientists are just beginning to investigate how the living brain performs the cognitive functions subsumed under the term *intelligence.*

Throughout the 20th century, however, psychologists have discovered many important facts about the nature of intelligent behavior itself. Some of the most well-established of these facts are realities of human nature that can be recognized and taken account of by our educational policies and practices to increase the benefits of education for every student in the school population. Ignoring these facts, I believe, limits the benefits of education to only some fraction of a nation's total population.

The most fundamental of these realities, as concerns public education, is the fact that there is an exceedingly wide range of individual differences in cognitive abilities in the school population. Indeed, the most troublesome concern of present-day universal public education seems to be the problem of how to deal with the conspicuous variation in ability. Before suggesting ways for the educational system to take account of this phenomenon successfully, I should spell out what we know of its most important aspects.

The Fields of Research on Mental Abilities

Ability here simply means a response that any organism can make to some form of stimulation, either external or internal, that is not an innate or hard-wired reflex and for which the organism's response can be objectively assessed, for example, in terms of its magnitude, latency (i.e., the time interval between stimulus and response), appropriateness (e.g., effective-in-effective, or correct-incorrect), and consistency. An ability is a *mental* ability only if individual differences in the ability are not mainly the result of individual differences in sensorimotor functions, such as visual and auditory acuity or physical strength and agility. This definition of *mental ability* is open-ended in that, for human beings at least, there is no theoretical limit to the number of different abilities that may exist, because there is no conceivable limit to the number of different things that persons may be able to do in the bounds of this definition of mental ability.

Behavioral research on human mental abilities has five major branches: (a) psychometrics, or the technology of measuring abilities; (b) differential psychology, or the study of the variety and correlational structure of individual and group (e.g., age, race, sex, social class) differences in abilities; (c) behavioral genetics, or study of the relative influences of genetic and environmental factors as causes of variation in abilities; (d) cognitive development, or the study of how abilities emerge, mature, differentiate, and develop throughout the individual's life span; and (e) instruction, or the applied psychology of inculcating knowledge and specific cognitive skills.

The first three of these branches are closely interrelated, but each one also constitutes a large and highly developed discipline in its own right;

together they constitute an important branch of behavioral science. Their aim is mainly scientific, that is, exploratory and explanatory, by proposing theories that suggest empirically testable hypotheses, performing studies and experiments that could disconfirm an hypothesis, and, if so, revising the theory. Much, or perhaps most, of the explanatory power that we now possess in this field, however, is of such a nature as to have little direct relevance to the main mission of education.

Research on mental ability has been mainly concerned with understanding the causes or sources of variance in many different behavioral measures and describing their covariance structure. *Variance* is simply a precise statistical measure of the degree of dispersion of individual differences in a particular measurable variable in a specified group of individuals or in a representative sample of some population. Covariance structure refers to the degree to which various measurements covary (or correlate) with one another. A particular variable, *X,* might have some part (or component) of its total variance in common with another variable, *Y,* for example, and both *X* and *Y* might also have components of variance in common with variable *Z.* If one and the same component of variance is common to all of the different variables (*X, Y,* and *Z*), it is called a *common factor.* The mathematical technique for analyzing a matrix of correlations among various measurements (for example, scores on a variety of tests) to determine the number of common factors and the proportion of the total variance accounted for by each factor is known as factor analysis.

One point of interest to the psychometrician is the number of independent common factors that exist among a large number of different tests that superficially appear to measure different mental abilities. Psychometricians are also interested in the relative "size" of each of the common factors, as indicated by the proportion of variance in all of the tests that is accounted for, or "explained" by, each factor.

The correlations among various measures of scholastic performance (e. g., spelling, reading comprehension, arithmetic, and knowledge of scholastic subject matter), along with measures of various nonscholastic mental abilities (such as various types of nonverbal IQ tests), can be analyzed together by means of factor analysis to discover their covariance structure. Finding the components of variance (i.e., individual differences) in scholastic performance and discovering their genetic and environmental sources of variance by the analytic methods of behavioral genetics are extremely different aims from those of research on the psychology of instruction, which tries to discover the best means for helping pupils attain maximum personal benefit from the time they spend in school. Yet, I believe educators need to know something about the results of factor analysis and behavior-genetic analysis to be more effective in the conduct of instruction throughout the whole educational process. So I try here to abstract from the research in psychometrics, differential psychology, and behavioral genetics what I consider the now well-established empirical findings that seem most relevant to the educational system's responsibility for dealing with the ubiquitous problem of individual differences in scholastic performance. These findings can be listed as a series of points. Then I indicate their implications for education and how they might be implemented in practice.

The Primacy of the G Factor in Scholastic Performance

The number of mental abilities (as defined earlier) is unlimited, but the number of independent common factors is relatively small. In other words, if we factor analyzed a great many different tests that were devised to measure as many seemingly different kinds of abilities as test constructors are able to imagine, we would find that, because of the great amount of correlation among the various tests, there are only a limited number of common factors represented in the whole lot. Without going into the technical details of factor analysis, it can be noted that the common factors differ greatly in their generality, that is, the number of different tests that reflect a particular factor and the amount of the tests' variance that is accounted for by a given factor. Factors that enter into only certain groups of tests based on a particular type of content are called *group factors*.

The more general group factors are labeled *verbal* (tests containing verbal material, such as vocabulary, reading comprehension, sentence completion, deciphering scrambled sentences), *spatial visualization* (tests based on non-verbal figural or puzzlelike problems, block designs, paper-folding, block counting), *numerical reasoning* (tests containing arithmetic problems, estimating quantities, number series), *mechanical reasoning* (tests based on problems involving the workings of balances, levers, cogwheels, gears, weights, and pulleys; see Carroll, 1993), and *memory* (tests requiring the recall of specific previously acquired information). There are also other factors of lesser generality, some of which may be called special talents, such as musical and artistic abilities (Carroll, 1993).

Over and above all of these group factors there exists one large superfactor that enters not only into tests of a particular type but into a great variety of tests of mental abilities, regardless of their specific knowledge content or required skills. There are few tests of mental ability that do not reflect this superfactor to some degree. A great deal of empirical evidence suggests that it may well be that the case that any test of mental ability, as defined earlier, reflects this superfactor to some degree, however slightly, when it is factor analyzed among a large and diverse collection of cognitive tests given to a large and representative sample of the general population (Carroll, 1993; Jensen, 1987, 1998).

This generalization is impossible to prove, because not every conceivable mental test can be tried. The generalization could be disproved, however, simply by discovering or constructing a test that qualified as a reliable measure of a mental ability under the foregoing definition, which, when factor analyzed under the conditions just specified, fails to show a statistically significant correlation with the superfactor, that is, a nonzero loading on the general factor of the highest order of the matrix of correlations among a large and diverse collection of mental tests. This has been attempted, but, to the best of my knowledge, no such demonstration has yet been accomplished (Alliger, 1988).

This ubiquitous superfactor is appropriately called *general mental ability.* It is a common source of variance across a great many cognitive tests of every description. Its existence was originally hypothesized in the mid-19th century, even before the invention of psychometric tests, but it was not discovered empirically until early in the 20th century, when, in 1904, the British psychologist Charles Spearman (1863-1945) invented a method of factor analysis that could prove the existence of a factor of general mental ability, which he labeled simply *g* (Spearman, 1927). The general factor in a battery or collection of conventional mental tests is typically referred to as *psychometric g.* It has proved to be one of the fundamental constructs of psychometrics and differential psychology.

Whatever is the ultimate cause of *g,* still the subject of ongoing research, its existence is manifested throughout positive correlations among diverse measures of cognitive ability. All positive correlations among various tests means that an individual's level of performance on any one test statistically predicts to some degree that individual's level of performance on any other test. Here are a number of things I think educators should know about the *g* factor.

Most important from the standpoint of education is the fact that *g* is the largest measurable source of common factor variance in scholastic performance, whether it is assessed by teacher's grades or by objective tests of scholastic achievement (Jensen, 1993a). In considering the entire school population, rather than just some restricted segment, it is a fact that *g* accounts for at least half of the total variance in scholastic achievement, which is far more than any other single source of variance independent of *g* (Cronbach & Snow, 1977; Gedye, 1981). The other measurable sources of variance that play second fiddle to *g* include all of the group factors of ability, special talents, specific non-*g* disabilities (e.g., dyslexia), personality variables, and students' socioeconomic status, sex, race, and national origin. By its very nature, the educational process, when it encompasses almost the entire population, highlights the effects of individual differences in level of *g* probably more than any other aspect of life in the modern industrialized world.

The strong relationship between individual differences in *g* and in occupational status and income in our society is largely mediated by the even stronger relationship between *g* and educational attainments (Jensen, 1993a; Snow & Yalow, 1982). We all take delight, of course, in stories of school failures or dropouts who go on to become billionaires or win Nobel Prizes, but these exceptional cases can by no means contradict the statistics based on large representative samples of a nation's population.

Although the *g* factor accounts for considerably more of the variance than any one of the group factors in a large number of diverse mental tests, it accounts for no more than about one third of the total variance in all of the wide variety of tests known to psychologists and that have been factor analyzed.

Unlike any specific ability or any group factor, *g* cannot be described in terms of the informational content or the types of skills called upon by the tests that are loaded with *g.* Vocabulary and block design tests, for example, are about equally *g*-loaded, yet they superficially have nothing in common. The block design test, for example, is a completely nonverbal test in which the participant is shown pictures of simple 2-dimensional designs and is asked to create a copy of the design using a set of colored blocks. The fact that *g* is loaded to varying degrees in every type of mental test implies that it cannot be described in terms of any one type of test.

Although all tests of mental ability are loaded on *g,* some tests have much larger *g* loadings than others; on a scale of 0 to 1, the *g* loadings of virtually all existing mental tests of any kind range continuously between about .20 and .90 (i.e., from 20% tp 90% of the total variance in different tests).

The degree of *complexity* (but not necessarily the difficulty) of the mental operations required by a test (or task of any kind) generally distinguishes between those tests with high *g* loadings and those with low *g* loadings (Jensen, 1987).

Not only formal psychometric tests but also many kinds of real-life activity that make cognitive demands, such as thinking, remembering, judging, decision making, reasoning, learning from experience, recalling relevant acquired knowledge or skills, and the like, are *g* loaded to varying degrees. Therefore *g* is reflected in a great many aspects of life, not just those generally thought of as scholastic, academic, or intellectual (Gordon, 1997; Gottfredson, 1997).

Individual differences in the best practical measures of *g,* such as the IQ obtained from a battery of diverse tests (e.g., the Stanford-Binet Intelligence Scale [Thorndike, Hagen, & Sattler, 1986] or the Wechsler Intelligence scales [Matarazzo, 1972], or the Kaufman Assessment Battery for Children [Kaufman & Kaufman, 1983) are approximately normally distributed in the population. These widely used test batteries are individually administered by a trained psychologist; each battery consists of a dozen or more diverse subtests designed to measure ability factors such as verbal, numerical reasoning, spatial reasoning, and memory; and about half of the subtests in each battery are performance tests, that is, they do not require reading or any overt use of language. In these tests' standardization sample, the scores, scaled as IQ, from the familiar bell curve, which means that most of the population is clustered around the middle (or average) of the distribution curve, with gradually decreasing percentages of the population showing increasingly below-average or above-average scores. On the IQ scale, with the average IQ set at 100 and the standard deviation set at 15, 50% of the population falls between IQs of 90 and 110, whereas 25% falls below an IQ of 90 and 25% falls above an IQ of 110.

The *g* construct does not reside in the tests or in test items, which merely serve as vehicles for measuring individual differences in the level of whatever causal factors are reflected by scores on highly *g*-loaded tests. Because tests having nothing superficially in common that would cause them to be correlated because they possess elements of knowledge and skills in common, it is reasonable to infer that the basic causal factors reflected by *g* are not intrinsic to the tests or to the methodology of factor analysis but are a property of the brain, specifically those neural processes involved in information processing and which cause differences between individuals in their speed and efficiency of information processing.

I have presented this view in detail elsewhere (Jensen, 1987, 1992, 1993b, 1997b). In brief, my view is that cognition is a form of information processing, and the *g* factor reflects individual differences in information processing as manifested in functions such as attending, selecting, searching, internalizing, deciding, discriminating, generalizing, learning, remembering, and using incoming and past-acquired information to solve problems and cope with the exigencies of the environment. These properties are obviously critical determinants of the individual's response to any educational experience, either formal or incidental. Other views of the nature of *g* based on less elemental processes than I have hypothesized emphasize metaprocesses, which are hypothetical constructs that serve an executive or controlling function in the deployment of the more elemental information processes, such as planning, selection of appropriate schemata for various kinds of problem solving, and monitoring, and evaluating one's own performance (Sternberg & Gardner, 1982). There are even more radically different views of the nature of human intelligence, but it is beyond the scope of this chapter to describe each of them. A fair and brief treatment of these can be found in an article by Wagner and Sternberg (1986) and more comprehensively in a recent and admirably balanced but entirely noncritical book by Gardner, Kornhaber, and Wake (1996). Although recognizing the existence of *g,* Howard Gardner (1993) minimized its importance and believed that human abilities can best be described in terms of at least seven different multiple intelligences, a view that appears to have gained greater acceptance among educators than among students of psychometrics and differential psychology.

Individual differences in the functions represented by *g* are strongly rooted in biology. This is demonstrated by two classes of evidence.

The first line of evidence is the fact that hereditary or genetic factors are involved in individual differences in test scores, as shown by the methods of quantitative genetic analysis applied to mental test data from various kinships, such as measuring the degrees of resemblance in IQ between identical twins reared apart in different environments as compared with identical twins reared together in the same family environment, or between genetically unrelated children reared together in the same family environment. These methods have been used to estimate statistically the proportion of genetic variance in a trait (based on individual differences in the population). This proportion of genetic variance is called the *heritability* of the trait (Plomin, 1990).

The heritability of IQ increases with age, going from about .40 in early childhood to about .80 in later maturity (McGue, Bouchard, Iacona, & Lykken, 1993; Fulker, Cherny, & Cardon, 1993). By adolescence, genetic factors contribute a larger part of the variance in IQ than is contributed by nongenetic or environmental factors. It is psychometric *g,* rather than other components of test score variance, that reflects most of the genetic variance (Thompson, Detterman, & Plomin, 1991).

The best single predictor of the heritability of individual differences in the scores on a given cognitive test is that test's *g* loading as determined when the test is entered into a factor analysis among a wide variety of other tests (Jensen, 1987; Pedersen, Plomin, & McClearn, 1994). It is noteworthy that, by late adolescence, the nongenetic or environmental component of variance in IQ, at least in industrialized societies, is not attributable to differences in family background (parent's occupation, socioeconomic status, number of books in the home, and the like) but consists of a multitude of small, largely random microenvironmental causes that make for differences among siblings reared in the same family (Jensen, 1997a; Rowe, 1994). Some of these nongenetic sources of variance are of a biological nature, both prenatal and postnatal, that are related to health, nutrition, and other physical variables that can affect brain development.

The second line of evidence is that certain purely physical measurements are correlated with individual differences in mental test scores and that the relative sizes of these correlations are best predicted by the tests' *g* loadings (Jensen, 1987, 1993b). The more *g*-loaded tests generally show the stronger correlations with the physical measurements. One such physical variable is brain size, or measurements related to brain size, such as head size, internal cranial capacity, and the total volume of the brain (Rushton & Ankney, 1996). Brain volume can be determined precisely by means of magnetic resonance imaging (MRI). It shows correlations with IQ averaging close to +.4. Another physical variable that is correlated with IQ (and with *g*) is the latency and amplitude of the average evoked potential, which measure the speed and strength of the "spike" of electrical activity that occurs in the brain in response to a single auditory or visual stimulus, such as a click or a flash of light (Deary & Caryl, 1993).

Yet another physical correlate of IQ (which also varies for different tests according to their *g* loadings), is the brain's glucose metabolic rate, as measured by the positron emission tomography (PET) scan, when the individual is engaged in an information processing task, such as taking a mental test or playing a video game (Haier, 1993).

The level of all mental abilities represented by *g* gradually increase with age from infancy to maturity, and, to the degree that they are *g*-loaded, they slowly decline with age in later maturity and, more rapidly, in old age. Brain size, reaction time, nerve conduction velocity, and brain metabolism all show a similar trajectory across the life span (Baltes & Kliegl, 1985; Birren, 1974; Cattell, 1971).

Individual differences in rates of learning are also correlated with *g* and are surprisingly large. Research in the armed forces, for example, shows that the fastest learners among recruits in the training programs are able to attain a satisfactory criterion of performance in the learning of various kinds of subject matter and technical skills some 5 to 10 times faster than the slowest learners. And on measures of *g* obtained in the public school population it has been found that the upper 10% of pupils at 7.5 years of age are equal to the lower 10% of pupils at 15.5 years of age in general knowledge and level of scholastic achievement. While some high school juniors and seniors are learning calculus, some others are still struggling with fractions and long division, and still others have simply given up completely on learning math. Individual differences in the ability to learn any specific cognitive skills, for example, skills and knowledge of the kind that schooling is intended to inculcate, are largely a matter of *g*. In fact, the general factor measured in cognitive learning tasks is the same *g* factor as found in psychometric tests (Jensen, 1989a; Snow & Lohman, 1989; Snow & Yalow, 1982).

To the extent that IQ tests measure specific learned cognitive skills, the test scores can be raised by special coaching or training in those specific skills. A gain in retest scores (or scores on highly similar tests) then largely reflects merely narrow transfer of training rather than an increased level of *g,* which is not a skill at all. I have found no bona fide evidence that the general ability represented by *g* is itself enhanced by such training (Jensen, 1993a). The prolonged and intensive effort experimentally to "train up" the level of *g* in children results in improved scores on IQ tests, but apparently what has been trained up in these cases is not *g,* but some of the specific learnable skills involved in conventional tests. Although the IQ scores of the tutored children are tested to be somewhat higher, these children, on average seldom perform significantly better scholastically than do age-matched control groups that were not so trained (Clarke & Clarke, 1976; Jensen, 1989b; Spitz, 1986). Moreover, the training effect rapidly fades within a year or so and the rates of mental growth of the groups of treated children and their scholastic progress merge with the same growth-curve trajectory as that of the untreated control group. In the range of human environments, the individual's growth curve for *g* is largely autonomous and hardly is deflected by external psychological or educational manipulations. This is probably because the main source of individual differences represented by *g* does not consist of a learned skill or set of specific skills but is a developmental aspect of the brain's structural capacity for information processing in general. In the Abecedarian Study (Ramey, 1992), one of the most methodologically rigorous interventions, disadvantaged children at high risk for low IQ and school failure were given intensive educational treatment from infancy to school age. At age 12, they showed a gain of 5.1 IQ points over an untreated control group (Ramey, 1994). At the last reported testing, when the children were 15 years of age, the difference between the treatment group and the control group was 4.6 IQ points (Campbell & Ramey 1994, 1995). The validity of even these modest findings and their interpretation, however, have been the subject of quite vehement debate (Ramey, 1993; Spitz, 1992, 1993a, 1993b).

As for social class and racial differences in *g,* the hypothesis that these group differences have the same causal basis as individual differences regardless of group membership is, I believe, at present most consistent with the preponderance of the empirical evidence related to this question (Rowe & Cleveland, 1997; Rowe, Vazsonyi, & Flannery, 1994, 1995). Mean differences among various population groups are viewed most reasonably simply as individual differences that have been aggregated in certain ways by a number of causal factors — evolutionary, cultural, religious, historical, political, and economic. Hence, I believe that the hypothesis for which there is the strongest evidence at present is that group differences in *g* (and its many correlates) are as real as individual differences, as variance in both individual and group differences are constituted of the same genetic and environmental sources of variance to about the same degree.

There are, of course, other views of the causes of racial and ethnic group differences, but those that I have examined either are contradicted by direct empirical evidence (Jensen, 1973, 1994), are philosophic opinions and historical speculation put forth in terms that would hardly be subject to empirically testable hypotheses (Gordon & Bhattacharyya, 1994), or are unsupported by relevant psychometric or statistical evidence (Ogbu, 1978; Sowell, 1994). The issue and its relevant evidence are complex, and involve a number of fields in the behavioral and biological sciences; therefore a proper discussion of it would be well beyond the scope of this chapter. Fairly comprehensive and scientifically respectable treatments of the topic are presented elsewhere (Herrnstein & Murray, 1994; Jensen, 1998; Rushton, 1995).

Mental abilities and other psychological traits range widely in every major population group. Therefore, identifying an individual in terms of race, sex, social class, or any other social group membership, affords no justifiable basis for differential educational prescription. This must depend on each pupil's own capabilities, not on group membership.

Finally, as important as *g* is in accounting for variance in scholastic performance, the sum total of educational attainments, occupational level, or career success, it is certainly not the only factor on which these outcomes depend (Shurkin, 1992). This becomes especially apparent when the life histories of the successful and the unsuccessful and of high achievers and ne'er-do-wells are contrasted in all these spheres of performance. Their *g* distributions markedly overlap. It is clear, however, that for any given type of achievement, depending on the degree of complexity of the demands it makes on information processing, the level of *g* required must exceed a certain threshold, or minimum level, to predict any realistic probability of success. This minimum level is inferred from the lowest level of IQ found in testing a large and representative sample of individuals in any given occupation. These minimum IQ levels for over 500 different occupations range from about 40 to 115 (Jensen, 1980,). For example, no one among persons employed as mathematicians was found by the U.S. Employment Service to have an IQ below 115. The probability, therefore, that anyone with an IQ truly less than 115 would become a professor of mathematics at Harvard or Berkeley, for example, is practically nil. Similarly, almost every occupation has its minimum IQ level, which is determined largely by the occupation's demands for information processing and knowledge acquisition. The probability of success in any given pursuit, to the extent that it makes such demands, increases with an increasing level of *g*.

But here is the important point: At every level of *g* above the requisite minimum threshold for a particular activity, the level of *g* itself is only a necessary but never a sufficient condition for success in any given educational or occupational pursuit. Other criteria for success, such as a successful marriage or a successful social life are much less related, if at all, to the individual's level of *g*. In fact, biographies of some of the greatest intellects in history, for example, Newton, Wagner, and Einstein, indicate that these geniuses were personally even less than mediocre in many respects, including what might be called social intelligence. A level of *g* greater than the threshold is only one of the variables among a number of other individual-difference variables necessary for outstanding performance, even in mainly intellectual pursuits. Personality variables, such as conscientiousness, dependability, emotional stability, energy level, motivation, zeal, and persistence of effort in the face of difficulty or hardship, as well as social skills and interpersonal sensitivity also play a part. For certain types of achievement, some special talent is an essential element in the equation (for example, musical, artistic, literary, mathematical, athletic, or aesthetic sensitivity). It is now believed that, in determining success, all these kinds of personal variables (including *g*) do not work in an additive way but rather in a multiplicative way, so that no one of the several variables needed for a particular type and level of achievement can be lacking (or be below some critical threshold). Thus an above-threshold level of any one of the several variables required for a particular accomplishment is said to be a necessary but not sufficient condition for achievement. This is as true for educational attainments as it is for occupational attainments.

On the strictly personal side in human relationships, of course, there are obviously many different reasons for one's appreciating and valuing a person regardless of his or her endowments in *g* or other traits associated with worldly achievements.

Educational Implications of Individual Differences in *g*

The American educational system is having trouble living with the consequences of *g,* which is essentially the problem of unyielding individual differences in the basic information processes on which educability depends (Jensen, 1991). The egalitarian notion that equal educational opportunity should lead to equal outcome, however, is of recent origin. The origins of schooling, in Europe (and later in America), were mainly concerned with educating an intellectual elite, to become the society's professional, managerial, and governing classes. The idea of universal public education came much later. Schools had traditionally expected many children to fail in school or to drop out when their rate of progress lagged markedly behind that of their peers in the lock-step advancement through the grades governed by pupils' chronological age. In the 1920s, for example, fewer than 30% of schoolchildren graduated from high school, as compared with more than 70% in 1960 (Herrnstein & Murray, 1994). Elementary school through 8th grade was in large part a means of screening pupils and selecting the academically talented (Chapman, 1988). This practice considerably reduced the problems associated with the increasingly wide range of individual differences in academic aptitude as students entered high school.

In the early industrial and still largely agrarian society of the 19th century, the most prevalent types of employment available to the majority of the population made comparatively small demands for scholastic skills and types of knowledge much beyond the rudiments of the "three Rs", which could be rather easily acquired by a majority of pupils in 5 or 6 years in the classroom. It was not expected, nor was it even essential, at the time, that a majority would or should continue on to more advanced levels of education. The lifetime consequences of failing school or dropping out at an early age were much less serious then than they are now in our technological and information-intensive society. The traditional system of education, unfortunately, has not changed nearly as radically as the society has changed in its technological employment demands for a more highly educated populace.

The traditional method of schooling is based on one teacher in a classroom of some 30 pupils, all at about the same chronological age, all progressing through standardized lessons and grade levels in a lock-step fashion. The real trouble began when the educational system insisted on universal school attendance from kindergarten through high school without fundamentally altering the conventional lock-step system, and then even tried to do away with ability grouping or tracking. The old system clearly is not set up to do the job that is needed to prepare the vast majority of the school population for productive participation in the nation's workforce in the years ahead.

The major changes that I consider necessary for updating the educational system and materially improving its effectiveness for the whole population, rather than for just the upper half of the bell curve, are so radical, as compared to what has been tried so far, that it cannot possibly consist merely of recommending practices to be carried out by the individual classroom teacher working in the constraints of the present methods of instruction. I suspect that most teachers are doing about as good a job as they possibly can under the present conditions. Any suggestions for them to do something a little different from what they are already doing, I think, will scarcely make a dent in the main problem of today's schools.

The root problem is that the schools are not taking adequate account of individual differences in *g,* and every segment of the bell curve is neglected as a result, especially the upper and the lower quartiles of the *g* distribution. Not until this shortcoming is dealt with effectively will achievements relevant to the employment needs of modern society be materially raised for students at every level of *g*.

It has been found, however, that successful attempts to raise the level of scholastic achievement with computer assisted instruction throughout every segment of the bell curve increases the overall mean level of achievement and individual levels of achievement, but also increases the spread of individual differences (Atkinson, 1974). Despite the inevitable increase in variance, which runs against egalitarian notions, because achievement gains can occur at every level of *g,* a larger number of students will come out above the critical thresholds on educational knowledge and skills required for gainful employment in our modern society. The net result is a positive gain, both for individuals and for society as a whole.

Besides the essential facts of *g* that I have outlined, the recognition of six fundamental principles govern suggestions for educational practice. They are based not on anyone's sentiments or what we might wish for but on what is actually known empirically:

1. *All* children can learn. And all healthy children initially want to learn and like to learn, unless their keenness for learning has been "turned off" by repeated failure and frustration.

2. Some children learn more quickly than others, and these differences in the general school population are greater than most people, including many educators, seem aware.

3. A child learns certain things faster and more easily than other things, and children differ in the things that, for them, are easier or harder to learn.

4. Individual differences in intrinsic motivation for cognitive learning appear to be at least as great as individual differences in *g* and are to some extent related to *g,* although we as yet have little understanding of the causes of differences in intrinsic motivation for cognitive activity.

5. Not every individual will attain a useful mastery of concepts and skills beyond a certain level of complexity and abstraction. That is, there are marked individual differences in the points at which different persons "top out" along the trajectory toward mastery of an intellectual subject matter or skill of a type that has virtually no ceiling on its most advanced level of complexity, regardless of the method of instruction and the amount of time and effort applied. In writing skill, mathematics, or musical composition, for example, not everyone can become a Shakespeare, a Newton, or a Beethoven. Most everyone who tries nevertheless "tops out" far below such towering levels of achievement, however strong may be their ambition and effort.

6. The best learning environment is like a good cafeteria. It not only affords the essential staples but also offers a large variety of choices to satisfy individual tastes. This allows children to discover their natural interests, proclivities, and special talents. It does not try to force round pegs into square holes. But pupils must also make some effort to learn enough about a variety of subjects and skills to enable the discovery of their areas of strongest potential for mastery. The motivation for such effort initially may have to be instigated and stimulated by parents and teachers. The father of the three Beethoven brothers had to make them practice their piano lessons; two of them discovered they were musical duds and quit taking music lessons; one of them, however, discovered he was a musical genius, but even he at first had to be made to practice, and, of course, the rest of his story is history. Psychologists now know that, at the higher levels of achievement, success is attained by individuals' trying and selecting those pursuits that best suit their particular aptitudes and interests.

In accord with these guiding principles, what sort of educational practices would probably be most effective for a system of universal public education designed to meet the future needs of our technological society?

In the most general terms, it would take into account ability differences, not by trying to eliminate variance, but by designing programs of educational input that can maximize the acquisition of the knowledge and skills most apt to benefit students (and society) after they leave school.

From the beginning of schooling, this would require a highly branching educational program, going from the same initial instruction in "basics" for all children, and thereafter branching frequently and extensively but never in a uniform or lock-step fashion for any particular group of children, whether by chronological age, or mental age, or by any other group classification. Progress through the instructional program would be geared entirely to each individual's performance at each point in time.

The kind of program I envision is impossible in the present-day classroom. If it were to be conducted under such conditions, it would require much more than a feasible extent of homogeneous ability grouping, much more testing than anyone would want, far too much continual shifting of children from one learning track to another, and way more instructional diversity than any one classroom teacher could possibly manage.

The model I have in mind can be likened more to private tutoring than to classroom instruction. It would take some 2 to 4 hours of the school day, for each pupil, with a detailed record kept of every phase of the pupil's progress, recorded during each tutorial session. Continual analyses of these records of a pupil's progress during the course of instruction, with periodic extraction of certain key summary information, would obviate any need for giving children tests as such. Measures of achievement would be a product of the learning program itself.

To accomplish this task entirely with school personnel, of course, would be wholly unfeasible. But we now have the technical capability of doing it with the aid of computer hardware and software. Further research and development is needed, however, to advance the computer software applications of the proposed technology beyond the present state of the art. Now that future educators are being recruited from the generation of youths who have grown up with computers, the time seems ripe for the developments I propose.

To use what now seems to me a rather old-fashioned term, I am referring to the development of computer-assisted instruction, (CAI), with computerized testing as a built-in and practically indistinguishable feature of the instructional program. The child's "summary key" of performance at any given point in time would be informative of the child's actual criterion-referenced achievement in the specific instructional program. It would also assess the extent of the generalizability, or breadth of transfer, of what had been learned, and would indicate the pupil's probable readiness for succeeding in the next step in the instructional program. Program branching to more advanced lessons or to remedial lessons would depend on whether certain criterion levels of performance were either attained or consistently not attained by varied instruction.

Pupils in the classroom would not be tracked or grouped in the early grades. Certain periods of the school day would be spent in small-group activities with the teacher. This is important because computers, of course, cannot be role models or potent sources of enthusiasm for learning a given subject. They cannot inculcate a love of learning, which good teachers are able to instill in their students. (In personal retrospect, I now realize that I got the most from those few teachers who evinced some genuine enthusiasm for the subject they were teaching, and in a few cases their effects on my interests and values are still evident some 50 or 60 years later; they were the few teachers who aroused in me what seemed at the time to be a spontaneous and autonomous interest in the subjects they taught.) One would not want a school that is so automated by computer-assisted instruction (CAI) that this kind of teacher influence would be lessened. Teachers' personal influence must work hand in hand with CAI. Also, in addition to instruction in academic subjects, pupils would need time to explore and develop their potential talents in writing, art, music, dance, and athletics.

I see no need in this picture for ability or achievement testing as such. The proper automatic recording and assessment of pupil performance during each instructional CAI period would supply all the information needed for the branching of the program and for the teacher's guidance of pupils. There would be no external gatekeepers who would exclude a student from entering a particular path for which he or she has demonstrated readiness through performance in prerequisites to the particular path. There would be no need for test scores per se or for tracking of pupils but only close monitoring of students' computer-recorded progress at every step of the way.

As pupils advance, the branching would become more and more extensive. Not all children would be expected to enter every branch or the same branches. Past performance would be the sole guide to future performance, and it would be tracked on a day-to-day basis. Backtracking and alternate branching at a later developmental age also would be possible. In math, for example, some students might branch into advanced algebra or calculus while some of their age peers topped out (at least temporarily) at, say, mixed fractions. They would branch into programs on applied math useful in common everyday activities — weighing and measuring things accurately, balancing a checkbook, verifying a bank statement, computing compound interest, and the like. Not measures of *g* per se, but actual performance and achievement level would be the basis for evaluation of the student's progress and advancement.

Also, by high-school age, ideally, real work apprenticeship programs would be available for students who wished to learn specific job skills that are best learned under supervision on the job. Many jobs, such as sailing a boat, cooking a dinner, or playing a musical instrument, require the automatization of many subskills and are not all that interesting or learnable without "hands on" experience and the informative feedback that comes only from trying to perform the task itself. Imagine trying to learn how to swim by listening to a lecture or reading a book on swimming!

A universal benefit of the CAI approach is that nearly all young people leaving school and entering this information-intensive and computer-dominated technological would be computer literate as if by second nature, because they would have worked with computers and learned through computers throughout their schooling. CAI would be a postgraduate mode of learning that would present no problems to those who must continue learning in order to advance in their jobs or keep up with new developments in their fields.

One can easily envision a large-scale industry growing up in the educational applications of computers. What is available to the schools in this line today, compared to what is technically feasible, given the development of a full-fledged educational research specialty in the design of instructional and evaluative computer software, could be like comparing the first airplanes with the modern jets.

Testing the efficacy of my proposal probably could be accomplished most effectively if it were instituted in schools with children whose age range from 5 to 18 years and who represent the typical range of individual differences. It would seem most feasible if this proposal were conducted under the auspices of a long-term university-based research project, in which a large team of specialists could coordinate their expertise in an all-out effort that would be the equivalent of the Human Genome Project for education.

References

Alliger, G.M. (1988). Do zero correlations really exist among measures of different intellectual abilities? *Educational and Psychological Measurement, 48,* 275-280.

Atkinson, R.C. (1974). Teaching children to read using a computer. *American Psychologist, 29,* 169-178.

Baltes, P.B., & Kliegl, R. (1985). On the dynamics between growth and decline in the aging of intelligence and memory. In K. Poeck, H-J Freund, & H. Gänshirt (Eds.), *Neurology* (pp. 1-17). Heidelberg: Springer-Verlag.

Birren, J.E. (1974). Translations in gerontology — from lab to life: Psychology and speed of response. *American Psychologist, 29,* 808-815.

Campbell, F.A., & Ramey, C.T. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. *Child Development, 65,* 684-698.

Campbell, F.A., & Ramey, C.T. (1995). Cognitive and school outcomes for high-risk African-American students at middle-adolescence: Positive effects of early intervention. *American Educational Research Journal, 32,* 743-772.

Carroll, J.B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge, England: Cambridge University Press.

Cattell, R.B. (1971). *Abilities: Their structure, growth, and action.* Boston: Houghton Mifflin.

Chapman, P.D. (1988). *Schools as sorters: Lewis M. Terman, applied psychology, and the intelligence testing movement, 1890-1930.* New York: N.Y. University Press.

Clarke, A.M., & Clarke, A.D.B. (1976). *Early experience: Myth and evidence.* New York: Free Press.

Cronbach, L.J., & Snow, R.E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions.* New York: Irvington.

Deary, I.J., & Caryl, P.G. (1993). Intelligence, EEG, and evoked potentials. In P.A. Vernon (Ed.), *Biological approaches to the study of human intelligence* (pp. 259-315). Norwood, NJ: Ablex.

Fulker, D.W., Cherny, S.S., & Cardon, L.R. (1993). Continuity and change in cognitive development. In R. Plomin & G.E.

McClearn (Eds.), *Nature, nurture, & psychology* (pp. 77-97). Washington, D.C.: American Psychological Association.

Gardner, H. (1993). *Frames of mind.* New York: Basic Books.

Gardner, H., Kornhaber, M.L., & Wake, W.K. (1996). *Intelligence: Multiple perspectives.* Fort Worth, TX: Harcourt Brace.

Gedye, C.A. (1981). *Longitudinal study (grades 1 through 10) of school achievement, self-confidence, and selected parental characteristics.* Doctoral dissertation, University of California, Berkeley.

Gordon, E.W., & Bhattacharyya, M. (1994). Race and intelligence. In R.J. Sternberg (Ed.), *Encyclopedia of human intelligence,* Vol. 2 (pp. 889-899). New York: Macmillan.

Gordon, R.A. (1997). Everyday life as an intelligence test: Effects of intelligence and intelligence context. *Intelligence, 24,* 203-320.

Gottfredson, L.S. (1997). Why *g* matters: The complexity of everyday life. *Intelligence, 24,* 79-132.

Haier, R.J. (1993). Cerebral glucose metabolism and intelligence. In P.A. Vernon (Ed.), *Biological approaches to the study of human intelligence* (pp. 317-373). Norwood, NJ: Ablex.

Herrnstein, R.J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life.* New York: Free Press.

Jensen, A.R. (1973). *Educability and group differences.* New York: Harper & Row.

Jensen, A.R. (1980). *Bias in mental testing.* New York: Free Press.

Jensen, A.R. (1987). The *g* beyond factor analysis. In R.R. Ronning, J.A. Glover, J.C. Conoley, & J.C. Witt (Eds.), *The influence of cognitive psychology on testing* (pp. 87-142). Hillsdale, NJ: Lawrence Erlbaum Associates.

Jensen, A.R. (1989a). The relationship between learning and intelligence. *Learning and Individual Differences, 1,* 37-62.

Jensen, A.R. (1989b). Raising IQ without increasing *g*? A review of *The Milwaukee Project: Preventing mental retardation in children at risk* by H.L. Garber. *Developmental Review, 9,* 234-258.

Jensen, A.R. (1991). Spearman's *g* and the problem of educational equality. *Oxford Review of Education, 17,* 169-187.

Jensen, A.R. (1992). Understanding *g* in terms of information processing. *Educational Psychology Review, 4,* 271-308.

Jensen, A.R. (1993a). Psychometric *g* and achievement. In B.R. Gifford (Ed.), *Policy perspectives on educational testing* (pp. 117-227). Boston: Kluwer Academic Publishers.

Jensen, A.R. (1993b). Spearman's *g*: Links between psychometrics and biology. *Annals of the New York Academy of Sciences, 702,* 103-131.

Jensen, A.R. (1994). Race and IQ scores. In R.J. Sternberg (Ed.), *Encyclopedia of human intelligence,* Vol. 2 (pp. 899-907). New York: Macmillan.

Jensen, A.R. (1997a). The puzzle of nongenetic variance. In R.J. Sternberg & E.L. Grigorenko (Eds.) *Intelligence, heredity and environment* (pp. 42-88). Cambridge: Cambridge University Press.

Jensen, A.R. (1997b). The neurophysiology of *g*. In C. Cooper & V. Varma (Eds.), *Processes in individual differences* (pp. 107-124). London: Rouledge.

Jensen, A.R. (1998). *The g factor.* Westport, CT: Praeger, Kaufman, A.S., & Kaufman, N.L. (1983). *K-ABC: Interpretive manual.* Circle Pines:, MN: American Guidance Service.

Kaufman, A.S., & Kaufman, N.L. (1983). *Kaufman assessment battery for children: Interpretive manual.* Circle Pines, MN: American Guidance Service.

Matarazzo, J.D. (1972). *Wechsler's measurement and appraisal of adult intelligence.* 5th ed. Baltimore: Williams & Wilkins.

McGue, M., Bouchard, T.J., Jr., Iacono, W.G., & Lykken, D.T. (1993). Behavioral genetics of cognitive ability: A life-span perspective. In R. Plomin & G.E. McClearn (Eds.), *Nature, nurture, and psychology* (pp. 59-76). Washington, D.C.: American Psychological Association.

Ogbu, J.U. (1978). *Minority education and caste: The American system in cross-cultural perspective.* New York: Academic Press.

Pedersen, N.L., Plomin, R. & McClearn, G.E. (1994). Is there G beyond *g*? (Is there genetic influence on specific cognitive abilities independent of genetic influence on general cognitive ability?). *Intelligence, 18,* 133-143.

Plomin, R. (1990). *Nature and nurture: An introduction to human behavioral genetics.* Pacific Grove, CA: Brooks/Cole.

Ramey, C.T. (1992). High-risk children and IQ: Altering intergenerational patterns. *Intelligence, 16,* 239-256.

Ramey, C.T. (1993). A rejoinder to Spitz's critique of the Abecedarian experiment. *Intelligence, 17,* 25-30.

Ramey, C.T. (1994). Abecedarian Project. In R.J. Sternberg (Ed.) *Encyclopedia of human intelligence* (Vol. 1, pp. 1-3). New York: Macmillan.

Rowe, D.C. (1994). *The limits of family influence: Genes, experience, and behavior.* New York: Guilford Press.

Rowe, D.C., & Cleveland, H.H. (1997). Academic achievement in African Americans and whites: Are the developmental processes similar? *Intelligence, 23,* 205-228.

Rowe, D.C., Vazsonyi, A.T., & Flannery, D.J. (1994). No more than skin deep: Ethnic and racial similarity in developmental process. *Psychological Review, 101,* 396-413.

Rowe, D.C., Vazsonyi, A.T., & Flannery, D.J. (1995). Ethnic and racial similarity in developmental process: A study of academic achievement. *Psychological Science, 6,* 33-38.

Rushton, J. P. (1995). *Race, evolution, and behavior: A life history perspective.* New Brunswick, NJ: Transaction.

Rushton, J.P., & Ankney, C.D. (1996). Brain size and cognitive ability: Correlations with age, sex, social class, and race. *Psychonomic Bulletin and Review, 3,* 21-36.

Shurkin, J.N. (1992). *Terman's kids.* Boston: Little, Brown & Co. Snow, R.E. & Lohman, D.F. (1989). Implications of cognitive psychology for educational measurement. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 263-331).

Snow, R.E., & Lohman, D.F. (1989). Implications of cognitive psychology for educational measurement. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 263-331. New York: Macmillan.

Snow, R.E. & Yalow, E. (1982). *Education and intelligence.* In R.J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 493-585). Cambridge: Cambridge University Press.

Sowell, T. (1994). *Race and culture: A world view.* New York: Basic Books.

Spearman, C. (1927). *The abilities of man: Their nature and measurement.* New York: Macmillan.

Spitz, H.H. (1986). *The raising of intelligence: A selected history of attempts to raise retarded intelligence.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Spitz, H.H. (1992). Does the Carolina Abecedarian early intervention project prevent sociocultural mental retardation? *Intelligence, 16,* 225-237.

Spitz, H.H. (1993a). When prophecy fails: On Ramey's response to Spitz's critique of the Abecedarian Project. *Intelligence, 17,* 17-23.

Spitz, H.H. (1993b). Spitz's reply to Ramey's response to Spitz'z first reply to Ramey's first response to Spitz's critique of the Abecedarian Project. *Intelligence, 17,* 31-35.

Sternberg, R.J., & Gardner, M.K. (1982). A componential interpretation of the general factor in human intelligence. In H.J. Eysenck (Ed.), *A model for intelligence* (pp. 231-254). New York: Springer-Verlag.

Thompson, L.A., Detterman, D.K., & Plomin, R. (1991). Associations between cognitive abilities and scholastic achievement: Genetic overlap but environmental differences. *Psychological Science, 2,* 158-165.

Thorndike, R.L., Hagen, E., & Sattler, J. (1986). *The Stanford-Binet Intelligence Scale: Fourth edition, Technical manual.* Chicago: Riverside.

Wagner, R.K., & Sternberg, R.J. (1986). Alternative conceptions of intelligence and their implications for education. In T.M. Tomlinson & H.J. Walberg (Eds.), *Academic work and educational excellence: Rasing student productivity* (pp. 223-283). Berkeley, CA: McCutchan.