How Much Can We Boost IQ and Scholastic Achievement?

ARTHUR R. JENSEN
University of California, Berkeley

Originally published in *Harvard Educational Review,* Vol. 39, No. 1, Winter 1969, pages 1-123.

Arthur Jensen argues that the failure of recent compensatory education efforts to produce lasting effects on children's IQ and achievement suggests that the premises on which these efforts have been based should be reexamined.

He begins by questioning a central notion upon which these and other educational programs have recently been based: that IQ differences are almost entirely a result of environmental differences and the cultural bias of IQ tests. After tracing the history of IQ tests, Jensen carefully defines the concept of IQ, pointing out that it appears as a common factor in all tests that have been devised thus far to tap higher mental processes.

Having defined the concept of intelligence and related it to other forms of mental ability, Jensen employs an analysis of variance model to explain how IQ can be separated into genetic and environmental components. He then discusses the concept of "heritability," a statistical tool for assessing the degree to which individual differeances in a trait like intelligence can be accounted for by genetic factors. He analyzes several lines of evidence which suggest that the heritability of intelligence is quite high (i.e., genetic factors are much more important than environmental factors in producing IQ differences).

After arguing that environmental factors are not nearly as important in determining IQ as are genetic factors, Jensen proceeds to analyze the environmental influences which may be most critical in determining IQ. He concludes that prenatal influences may well contribute the largest environmental influence on IQ. He then discusses evidence which suggests that social class and racial variations in intelligence cannot be accounted for by differences in environment but must be attributed partially to genetic differences.

After he has discussed the influence on the distribution of IQ in a society on its functioning, Jensen examines in detail the results of educational programs for young children, and finds that the changes in IQ produced by these programs are generally small. A basic conclusion of Jensen's discussion of the influence of environment on IQ is that environment acts as a "threshold variable." Extreme environmental deprivation can keep the child from performing up to his genetic potential, but an enriched educational program cannot push the child above that potential.

Finally, Jensen examines other mental abilities that might be capitalized on in an educational program, discussing recent findings on diverse patterns of mental abilities between ethnic groups and his own studies of associative learning abilities that are independent of social class. He concludes that educational attempts to boost IQ have been misdirected and that the educational process should focus on teaching much more specfic skills. He argues that this will be accomplished most effectively if educational methods are developed which are based on other mental abilities besides I.Q.

The Failure of Compensatory Education

Compensatory education has been tried and it apparently has failed.

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

The chief goal of compensatory education—to remedy the educational lag of disadvantaged children and thereby narrow the achievement gap between "minority" and "majority" pupils—has been utterly unrealized in any of the large compensatory education programs that have been evaluated so far. On the basis of a nationwide survey and evaluation of compensatory education programs, the United States Commission on Civil Rights (1967) came to the following conclusion:

> The Commission's analysis does not suggest that compensatory education is incapable of remedying the effects of poverty on the academic achievement of individual children. There is little question that school programs involving expenditures for cultural enrichment, better teaching, and other needed educational services can be helpful to disadvantaged children. The fact remains, however, that none of the programs appear to have raised significantly the achievement of participating pupils, as a group, within the period evaluated by the commission (p. 138).

The commission's review gave special attention to compensatory education in majority-Negro schools whose programs "were among the most prominent and included some that have served as models for others." The Commission states: "A principal objective of each was to raise the academic achievement of disadvantaged children. Judged by this standard the programs did not show evidence of much success (p. 138).[1]

Why has there been such uniform failure of compensatory programs wherever they have been tried? What has gone wrong? In other fields, when bridges do not stand, when aircraft do not fly, when machines do not work, when treatments do not cure, despite all conscientious efforts on the part of many persons to make them do so, one begins to question the basic assumptions, principles, theories, and hypotheses that guide one's efforts. Is it time to follow suit in education?

The theory that has guided most of these compensatory education programs, sometimes explicitly, sometimes implicitly, has been two main complementary facets: one might be called the "average children concept," the other the "social deprivation hypothesis."

The "average children" concept is essentially the belief that all children, except for a rare few born with severe neurological defects, are basically very much alike in their mental development and capabilities, and that their apparent differences in these characteristics as manifested in school are due to rather superficial differences in children's upbringing at home, their preschool and out-of-school experiences, motivations and interests, and the educational influences of their family background. All children are viewed as basically more or less homogeneous, but are seen to differ in school performance because when they are out of school they learn or fail to learn certain things that may either help them or hinder them in their school work. If all children could be treated more alike early enough, long before they come to school, then they could all learn from the teacher's instruction at about the same pace and would all achieve at much the same level, presumably at the "average" or above on the usual grade norms.

The "social deprivation hypothesis" is the allied belief that those children of ethnic minorities and the economically poor who achieve "below average" in school do so mainly because they begin school lacking certain crucial experiences which are prerequisites for school learning—perceptual, attentional, and verbal skills, as well as the self-confidence, self-direction, and teacher-oriented attitudes conductive to achievement in the classroom. And they lack the parental help and encouragement needed to promote academic achievement throughout their schooling. The chief aim of preschool and compensatory programs, therefore, is to make up for these environmental lacks as quickly and intensively as possible by providing the assumedly appropriate experiences, cultural enrichment, and training in basic skills of the kind presumably possessed by middle-class "majority" children of the same age.

The success of the effort is usually assessed in one or both of two ways: by gains in IQ and in scholastic achievement. The common emphasis on gains in IQ is probably attributable to the fact that it can be more efficiently "measured" than scholastic achievement, especially if there is no specific "achievement" to begin with. The IQ test can be used at the very beginning of Headstart, kindergarten, or first grade as a "pre-test" against which to assess "post-test" gains. IQ gains, if they occur at all, usually occur rapidly, while achievement is a long-term affair. And probably most important, the IQ is commonly interpreted as indicative of a more general kind of intellectual ability than is reflected by the acquisition of specific scholastic knowledge and skills. Since the IQ is known to predict scholastic performance better than any other single measurable attribute of the child, it is believed, whether rightly or wrongly, that if the child's IQ can be appreciably raised, academic achievement by and large will take care of itself, given normal motivation and standard instruction. Children with average or above-average IQs generally do well in school without much special attention. So the remedy deemed logical for children who do poorly in school is to boost their IQs up to where they can perform like the majority—in short to make them all at least "average children." Stated so bluntly, the remedy may sound rather grim, but this is in fact essentially what we are attempting in our special programs of preschool enrichment and compensatory education. This simple theme, with only slight embellishments, can be found repeated over and over again in the vast recent literature on the psychology and education of children called culturally disadvantaged.

So here is where our diagnosis should begin—with the concept of the IQ: how it came to be what it is; what it "really" is; what makes it vary from one individual to another; what can change it, and by what amount.

The Nature of Intelligence

The nature of intelligence is one of the vast topics in psychology. It would be quite impossible to attempt to review here the main theoretical issues and currents of thought in this field. Large volumes have been written on the subject (e.g., Guilford, 1967; Stoddard, 1943), to say nothing of the countless articles. An enlightening brief account of the history of the concept of intelligence has been presented by Sir Cyril Burt (1968). The term "intelligence," as used by psychologists, is itself of fairly recent origin. Having been introduced as a technical term in psychology near the turn of the century, it has since filtered down into common parlance, and therefore some restriction and clarification of the term as it will be sued in the following discussion is called for.

Disagreements and arguments can perhaps be forestalled if we take an operational stance. First of all, this means that probably the most important fact about intelligence is that we can measure it. Intelligence, like electricity, is easier to measure than to define. And if the measurements bear some systematic relationships to other data, it means we can make meaningful statements bout the phenomenon we are measuring. There is no point in arguing the question to which there is no answer, the question of what intelligence *really* is. The best we can do is to obtain measurements of certain kinds of behavior and look at their relationships to other phenomena and see if these relationships make any kind of sense and order. It is from these orderly relationships that we can gain some understanding of the phenomena.

But how did the instruments by which we measure intelligence come about in the first place? The first really useful test of intelligence and the progenitor of nearly all present-day intelligence tests was the Metrical Scale of Intelligence devised in 1905 by Binet and Simon. A fact of great but often unrealized implications is that the Binet-Simon test was commissioned by the Minister of Public Instruction in Paris for the explicit purpose of identifying children who were likely to fail in school. It was decided they should be placed in special schools or classes before losing too much ground or receiving too much discouragement. To the credit of Binet and Simon, the test served this purpose quite well, and it is now regarded as one of the major "breakthroughs" in the history of psychology. Numerous earlier attempts to devise intelligence tests were much less successful from a practical standpoint, mainly because the kinds of functions tested were decided upon in terms of early theoretical notions about the basic elements of "mind" and the "brass instrument" laboratory techniques for measuring these elemental functions of consciousness, which were then thought to consist of the capacity for making fine sensory discriminations in the various sensory modalities. Although these measurements were sufficiently reliable, they bore little relationship to any "real life" or "common sense" criteria of behavior ranging along a "dull"-"bright" continuum. The psychological sagacity of Binet and Simon as test constructors derived largely from their intimate knowledge and observation of the behavior of young children and of what, precisely, teachers expected of them in school. Binet and Simon noted the characteristics distinguishing those children described by their teachers as "bright" from those described as "dull," and, from these observations and considerable trial-and-error, they were finally able to make up a graded series of test items that not only agreed with teachers' judgments of children's scholastic capabilities but could make the discriminations more finely and more accurately than any single teacher could do without prolonged observation of the child in class. The Binet-Simon scale has since undergone many revisions and improvements, and today, in the form developed by Terman, known as the Stanford-Binet Intelligence Scale, it is generally regarded as the standard for the measurement of intelligence.

But the important point I wish to emphasize here is that these Binet tests, and in effect all their descendants, had their origin in the educational setting of the Paris schools of 1900, and the various modifications and refinements they have undergone since then have been implicitly shaped by the educational traditions of Europe and North America. The content and methods of instruction represented in this tradition, it should be remembered, are a rather narrow and select sample of all the various forms of human learning and of the ways of imparting knowledge and skills. The instructional methods of the traditional classroom were not invented all in one stroke, but evolved within an upper-class segment of the European population, and thus were naturally shaped by the capacities, culture, and needs of those children whom the schools were primarily intended to serve. At least implicit in the system as it originally developed was the expectation that not all children would succeed. These methods of schooling have remained essentially unchanged for many generations. We have accepted traditional instruction so completely that it is extremely difficult even to imagine, much less to put into practice, any radically different forms that the education of children could take. Our thinking almost always takes for granted such features as beginning formal instruction at the same age for all children (universally between ages five and six), instruction of children in groups, keeping the same groups together in lock step fashion through the first several years of schooling, and an active-passive, showing-seeing, telling-listening relationship betwen teacher and pupils. Satisfactory learning occurs under these conditions only when children come to school with certain prerequisite abilities and skills: an attention span long enough to encompass the teacher's utterances and demonstrations, the ability voluntarily to focus one's attention where it is called for, the ability to comprehend verbal utterances and to grasp relationships between things and their symbolic representations, the ability to inhibit large-muscle activity and engage in covert "mental" activity, to repeat instruction to oneself, to persist in a task until a self-determined standard is attained—in short, the ability to engage in what might be called self-instructional activities, without which group instruction alone remains ineffectual.

The interesting fact is that, despite all the criticisms that can easily be leveled at the educational system, the traditional forms of instruction have actually worked quite well for the majority of children. And the tests that were specifically devised to distinguish those children least apt to succeed in this system have also proved to do their job quite well. The Stanford-Binet and similar intelligence tests predict various measures of scholastic achievement with an average validity coefficient of about .5 to .6, and in longitudinal data comprising intelligence tests and achievement measures on the same children over a number of years, the multiple correlation between intelligence and scholastic achievement is almost as high as the reliability of the measures will permit.

The Generality and Limitations of Intelligence

If the content and instructional techniques of education had been markedly different from what they were in the beginning and, for the most part, continue to be, it is very likely that the instruments we call intelligence tests would also have assumed a quite different character. They might have developed in such a way as to measure a quite different constellation of abilities, and out conception of the nature of intelligence, assuming we still called it by that name, would be correspondingly different. This is why I think it so important to draw attention to the origins of intelligence testing.

But in granting that the measurement and operational definitions of intelligence had their origins in a school setting and were intended primarily for scholastic purposes, one should not assume that intelligence tests measure *only* school learning or cultural advantages making for scholastic success and fail to tap anything of fundamental psychological importance. The notion is sometimes expressed that psychologists have mis-aimed with their intelligence tests. Although the tests may predict scholastic performance, it is said, they do not *really* measure intelligence—as if somehow the "real thing" has eluded measurement and perhaps always will. But this is a misconception. We *can* measure intelligence. As the late Professor Edwin G. Boring pointed out, intelligence, by definition, is what intelligence tests measure. The trouble comes only when we attribute more to "intelligence" and to our measurements of it than do the psychologists who use the concept in its proper sense.

The idea of intelligence has justifiably grown considerably beyond its scholastic connotations. Techniques of measurement not at all resembling the tasks of the Binet scale and in no way devised with the idea of predicting scholastic performance can also measure approximately the same intelligence as measured by the Binet scale. The English psychologist Spearman devoted most of his distinguished career to studying the important finding that almost any and every test involving any kind of complex mental activity correlates positively and substantially with any and every other test involving complex mental activity, regardless of the specific content or sensory modality of the test. Spearman noted that if the tests called for the operation of "higher mental processes," as opposed to sheer sensory acuity, reflex behavior, or the execution of established habits, they showed positive intercorrelations, although the tests bore no superficial resemblance to one another. They might consist of abstract figures involving various spatial relationships, or numerical problems, or vocabulary, or verbal analogies. For example, a vocabulary test shows correlations in the range of .50 to .60 with a test that consists of copying sets of designed with colored blocks; and a test of general information corelates about .50 with a test that involves wending through a printed maze with a pencil. Countless examples of such positive correlations between seemingly quite different tests can be found in the literature on psychological tests. Spearman made them the main object of his study. To account for the intercorrelations of "mental" tests, he hypothesized the existence of a single factor common to all tests involving complex mental processes. All such tests measure this common factor to some degree, which accounts for the intercorrelations among all the tests. Spearman called the common factor "general intelligence" or simply *g*. And he invented the method known as factor analysis to determine the amount of *g* in any particular test. He and his students later developed tests, like Raven's Progressive Matrices and Cattell's Culture Fair Tests of *g,* which measure *g* in nearly pure form. We should not reify *g* as an entity, of course, since it is only a hypothetical construct intended to explain covariation among tests. It is a hypothetical source of variance (individual differences) in test scores. It can be regarded as the nuclear operational definition of intelligence, and when the term intelligence is used it should refer to *g,* the factor common to all tests of complex problem solving.

In examining those tests most heavily loaded with *g,* Spearman characterized the mental processes which they seemed to involved as "the ability to educe relations and correlates"—that is, to be able to see the general from the particular and the particular as an instance of the general. A similar definition of intelligence was expressed by Aquinas, as "the ability to combine and separate"—to see the difference between things which seem similar and to see the similarities between things which seem different. These are essentially the processes of abstraction and conceptualization. Tasks which call for problem solving requiring these processes are usually the best measures of *g*. Despite numerous theoretical attacks on Spearman's basic notion of a general factor, *g* has stood like a rock of Gibraltar in psychometrics, defying any attempt to construct a test of complex problem solving which excludes it.

Standard intelligence scales such as the Binet and the Wechsler are composed of a dozen or so subtests which differ obviously in their superficial appearance: vocabulary, general information, memory span for digits, block designs, figure copying, mazes, form boards, and so on. When the intercorrelations among a dozen or more such tests are subjected to a factor analysis or principal components analysis, some 50 percent or more of the total individual differences variance in all the tests is usually found to be attributable to a general factor common to all the tests. Thus, when we speak of intelligence it is this general factor, rather than any single test, that we should keep in mind.

Attempts to assess age differences in intelligence or mental development which rely on complex techniques that bear little formal resemblance to the usual intelligence tests still manage to measure *g* more than anything else. Piaget's techniques for studying mental growth, for example, are based largely on the child's development of the concepts of invariance and conservation of certain properties—number, area, and volume. When a large variety of Piaget tasks are factor analyzed along with standard psychometric tests, including the Stanford-Binet and Raven's Progressive Matrices, it is found that the Piaget tasks are loaded on the general factor to about the same extent as the psychometric tests (Vernon, 1965). That is to say, children fall into much the same rank order of ability on all these cognitive tests. Tuddenham (1968) has developed a psychometric scale of intelligence based entirely upon Piaget's theory of cognitive development. The test makes use of ten of the techniques developed by Piaget for studying conservation, seriation, reversal of perspective, and so on. Performance on these tasks shows about the same relationship to social class and race differences as is generally found with the Stanford-Binet and Wechsler scales. It seems evident that what we call general intelligence can be manifested in many different forms and thus permits measurement by a wide variety of techniques. The common feature of all such intercorrelated tests seems to be their requirement of some form of "reasoning" on the part of the subject—some active, but usually covert, transformation or manipulation of the "input" (the problem) in order to arrive at the "output" (the answer).

The conceptually most pure and simple instance of this key aspect of intelligence is displayed in the phenomenon known as cross-modal transfer. This occurs when a person to whom some particular stimulus is exposed in one sensory modality can then recognize the same stimulus (or its essential features) in a different sensory modality. For example, show a person a number of differently shaped wooden blocks, then point to one, blindfold the person, shuffle the blocks, and let the person find the indicated block by using his sense of touch. Or "write" in bold strokes any letter of the alphabet between a child's shoulder blades. It will be a completely unique stimulus input for the child, never encountered before and never directly conditioned to any verbal response. Yet, most children, provided they already know the alphabet, will be able to name the letter. There are no direct neural connections between the visual and the tactile impressions of the stimulus, and, although the child's naming of the letter has been conditioned to the visual stimulus, the tactile stimulus has been associated with neither the visual stimulus nor the verbal response. How does the child manage to show the cross modal transfer? Some central symbolic or "cognitive" processing mechanism is involved, which can abstract and compare properties of "new" experiences with "old" experiences and thereby invest the "new" with meaning and relevance. Intelligence is essentially characterized by this process.

Is g Unitary or Divisible?

It is only when the concept of *g* is attributed meaning above and beyond that derived from the factor analytic procedures from which it gains its strict technical meaning that we run into the needless argument over whether *g* is a unitary ability or a conglomerate of many abilities. We should think of *g* as a "source" of individual differences in scores which is common to a number of different tests. As the tests change, the nature of *g* will also change, and a test which is loaded, say, .50 on *g* when factor analyzed, among one set of tests may have a loading of .20 or .80, or some other value, when factor analyzed among other sets of tests. Also, a test which, in one factor analysis, measures only *g* and nothing else, may show that it measures *g* and one or more other factors when factor analyzed in connection with a new set of tests. In other words, *g* gains its meaning from the tests which have it in common. Furthermore, no matter how simple or "unitary" a test may appear to be, it is almost always possible to further fractionate the individual differences variance to smaller subfactors. I have been doing this in my laboratory with respect to a very simple and seemingly "unitary" ability, namely, digit span (Jensen, 1967b). Changing the rate of digit presentation changes the rank order of subjects in their ability to recall the digits. So, too, does interposing a 10-second delay between presentation and recall, and interpolating various distractions ("retroactive inhibition") between presentation and recall, and many other procedural variations of the digit span paradigm. Many—but, significantly, not all—of these kinds of manipulations introduce new dimensions or factors of individual differences. It is likely that when we finally get down to the irreducible "atoms" of memory span ability, so to speak, if we ever do get there, the elements that make up memory span ability will not themselves even resemble what we think of as abilities in the usual sense of the term. And so probably the same would be true not only for digit span, but for any of the subtests or items that make up intelligence tests.

A simple analogy in the physical realm may help to make this clear. If we are interested in measuring general athletic ability, we can devise a test consisting of running, ball throwing, batting, jumping, weight lifting, and so on. We can obtain a "score" on each one of these and the total for any individual is his "general athletic ability" score. This score would correspond to the general intelligence score yielded by tests like the Stanford-Binet and the Wechsler scales.

Or we can go a step further in the refinement of our test procedure and intercorrelate the scores on all these physical tasks, factor analyze the intercorrelations, and examine the general factor, if indeed there is one. Assuming there is, we would call it "general athletic ability." It would mean that on all of the tasks, persons who excelled on one also tended to be superior on the others. And we would not that some tasks were more "loaded" with this general factor than others. We could then weight the subtest scores in proportion to their loading on *g* and then add them up. The total, in effect, is a "factor score," and gives us a somewhat more justifiable measure of "general athletic ability," since it represents the one source of variation that all the athletic skills in our test battery share in common.

To go still further, let us imagine that the running test has the highest loading on *g* in this analysis. To make the issue clear-cut, let us say that all its variance is attributable to the *g* factor. Does this mean that running ability is not further analyzable into other components? *No, it simply means that the components into which running can be analyzed are not separately or independently manifested in either the running test or the other tests in the battery.* But we can measure these components of running ability independently, if we wish to: total leg length, the ratio of upper to lower leg length, strength of leg muscles, physical endurance, "wind" or vital capacity, ratio of body height to weight, degree of mesomorphic body build, specific skills such as starting speed—all are positively correlated with running speed. And if we intercorrelate these measures and factor analyze the correlations, we would probably find a substantial general factor common to all these physical attributes, name it what you will. We could combine the measures on these various physical traits into a weighted composite score which would predict running ability as measured by the time the person takes to cross the finish line. The situation seems very similar to the analysis of the psychological processes that make up "general intelligence."

Fluid and Crystallized Intelligence

Raymond B. Cattell (1963) has made a conceptually valid distinction between two aspects of intelligence, *fluid* and *crystallized.* Standard intelligence tests generally measure both the fluid and crystallized components of *g*, and, since the two are usually highly correlated in a populations whose members to a large extent share a common background of experience, culture, and education, the fluid and crystallized components may not always be clearly discernible as distinct factors. Conceptually, however, the distinction is useful and can be supported empirically under certain conditions. *Fluid* intelligence is the capacity for new conceptual learning and problem solving, a general "brightness" and adaptability, relatively independent of education and experience, which can be invested in the particular opportunities for learning encountered by the individual in accord with his motivations and interests. Tests that measure mostly fluid intelligence are those that minimize cultural and scholastic content. Cattell's Culture Fair Tests and Raven's Progressive Matrices are good examples. *Crystallized* intelligence, in contrast, is a precipitate out of experience, consisting of acquired knowledge and developed intellectual skills. Fluid and crystallized intelligence are naturally correlated in a population sharing a common culture, because the acquisition of knowledge and skills in the first place depends upon fluid intelligence. While fluid intelligence attains its maximum level in the late teens and may even begin to decline gradually shortly thereafter, crystallized intelligence continues to increase gradually with the individual's learning and experience all the way up to old age.

Occupational Correlates of Intelligence

Intelligence, as we are using the term, has relevance considerably beyond the scholastic setting. This is so partly because there is an intimate relationship between a society's occupational structure and its educational system. Whether we like it or not, the educational system is one of society's most powerful mechanisms for sorting out children to assume different roles in the occupational hierarchy.

The evidence for a hierarchy of occupational prestige and desirability is unambiguous. Let us consider three sets of numbers.[2] First, the Barr scale of occupations, devised in the early 1920s, provides one set of data. List of 120 representative occupations, each definitely and concretely described, were give to 30 psychological judges who were asked to rate the occupations of a scale from 0 to 100 according to the grade of intelligence each occupation was believed to require for ordinary success. Second, in 1964, the National Opinion Research Center (NORC), by taking a large public opinion poll, obtained ratings of the *prestige* of a great number of occupations; these prestige ratings represent the average standing of each occupation relative to all the others in the eyes of the general public. Third, a rating of socioeconomic status (SES) is provided by the *1960 Census of Population: Classified Index of Occupations and Industries,* which assigns to each of the hundred listed occupations a score ranging from 0 to 96 as a composite index of the average income and educational level prevailing in the occupation.

The interesting point is the set of correlations among these three independently derived occupational ratings.

The Barr scale and the NORC ratings are correlated .91.

The Barr scale and the SES index are correlated .81.

The NORC ratings and the SES index are correlated .90.

In other words, psychologists' concept of the "intelligence demands" of an occupation (Barr scale) is very much like the general public's concept of the prestige or "social standing" of an occupation (NORC ratings), and both are closely related to an independent measure of the educational and economic status of the person pursuing an occupation (SES index). As O.D. Duncan (1968, pp. 90-91) concludes, "...'intelligence' is a socially defined quality and this social definition is not essentially different from that of achievement or status in the occupational sphere.... When psychologists came to propose operational counterparts to the notion of intelligence, or to devise measures thereof, they wittingly or unwittingly looked for indicators of capability to function in the system of key roles in the society." Duncan goes on to note, "Our argument tends to imply the a correlation between IQ and occupational achievement was more or less built into IQ tests, by virtue of the psychologists' implicit acceptance of the social standards of the general populace. Had the first IQ tests been devised in a hunting culture, 'general intelligence' might well have turned out to involve visual acuity and running speed, rather than vocabulary and symbol manipulation. As it was, the concept of intelligence arose in a society where high status accrued to occupations involving the latter in large measure, so that what we now *mean* by intelligence is something like the probability of acceptable performance (given the opportunity) in occupations varying in social status."

So we see that the prestige hierarchy of occupations is a reliable objective reality in our society. To this should be added the fact that there is undoubtedly some relationship between the levels of the hierarchy and the occupations' intrinsic interest, desirability, or gratification to the individuals engaged in them. Even if all occupations paid alike and received equal respect and acclaim, some occupations would still be viewed as more desirable than others, which would make for competition, selection, and, again, a kind of prestige hierarchy. Most persons would agree that painting pictures is more satisfying than painting barns, and conducting a symphony orchestra is more exciting than directing traffic. We have to face it: the assortment of persons into occupational roles simply is not "fair" in any absolute sense. The best we can ever hope for is that true merit, given equality of opportunity, act as the basis for the natural sorting process.

Correlation Between Intelligence and Occupational Achievement

Because intelligence is only one of a number of qualities making for merit in any given occupation, and since most occupations will tolerate a considerable range of abilities and criteria of passable performance, it would be surprising to find a very high correlation between occupational level and IQ. Although the rank order of the *mean* IQs of occupational groups is about as highly correlated with the occupations' standing on the three "prestige" ratings mentioned above as the ratings are correlated among themselves, there is a considerable dispersion of IQs *within* occupations. The IQ spread increases as one moves down the scale from more to less skilled occupations. (Tyler, 1965, pp. 338-339). Thus, the correlation, for example, between scores on the Army General Classification Test, a kind of general intelligence test, and status ratings of the civilian occupations of 18,782 white enlisted men in World War II was only .42. Since these were mostly young men, many of whom had not yet completed their education or established their career lines, the correlation of .42 is lower than one would expect in the civilian population. Data obtained by the U.S. Employment Service in a civilian population shows a correlation of .55 between intelligence and occupational status, a value which, not surprisingly, is close to the average correlation between intelligence and scholastic achievement (Duncan, et al., 1968, pp. 98-101). Although these figures are based on the largest samples reported in the literature and are therefore probably the most reliable statistics, they are not as high as the correlations found in some other studies. Two studies found, for example, that IQs of school boys correlated .57 and .71 with their occupational status 14 and 19 years later, respectively (Tyler, 1965, p. 343). It is noteworthy that the longer interval showed the higher correlation.

Duncan's (1968) detailed analysis of the nature of the relationship between intelligence and occupational status led him to the conclusion that "the bulk of the influence of intelligence on occupation is indirect, via education." If the correlation of intelligence with education and of education with occupation is, in effect, "partialled out," the remaining "direct" correlation between intelligence and occupation is almost negligible. But Duncan points out that this same type of analysis (technically known as "path coefficients analysis") also reveals the interesting and significant finding that intelligence plays a relatively important part as a cause of differential *earnings.* Duncan concludes: "...men with the same schooling and in the same line of work are differentially rewarded in terms of mental ability" (1968, p. 118).

Correlations Between Intelligence and Job Performance Within Occupations

Intelligence, via education, has its greatest effect in the assorting of individuals into occupational roles. Once they are in those roles, the importance of intelligence per se is less marked. Ghiselli (1955) found that intelligence tests correlate on the average in the range of .20 to .25 with ratings of actual proficiency on the job. The speed and ease of training for various occupational skills, however, show correlations with intelligence averaging about .50, which is four to five times the predictive power that the same tests have in relation to work proficiency *after* training. This means that, once the training hurdle has been surmounted, many factors besides intelligence are largely involved in success on the job. This is an important fact to keep in mind at later points in this article.

Is Intelligence Fixed?

Since the publication of J. McV. Hunt's well-known and influential book, *Intelligence and Experience* (1961), the notion of "fixed intelligence" has assumed the status of a popular cliché among many speakers and writers on intelligence, mental retardation, cultural disadvantage, and the like, who state, often with an evident sense of virtue and relief, that modern psychology has overthrown the "belief in fixed intelligence." This particular bugaboo seems to have loomed up largely in the imaginations of those who find such great satisfaction in the idea that "fixed intelligence" has been demolished once and for all.

Actually, there has been nothing much to demolish. When we look behind the rather misleading term "fixed intelligence," what we find are principally two real and separate issues, each calling for empirical study rather than moral philosophizing. Both issues lend themselves to empirical investigation and have long been subjects of intensive study. The first issue concerns the genetic basis of individual differences in intelligence; the second concerns the stability or constancy of the IQ throughout the individual's lifetime.

Genotype and Phenotype. Geneticists have avoided confusion and polemics about the issue of whether or not a given trait is "fixed" by asking the right question in the first place: how much of the variation (i.e., individual differences) in a particular trait or characteristic that we observe or measure (i.e., the *phenotype*) in a given population can we account for in terms of variation in the genetic factors (i.e., the *genotype*) affecting the development of the characteristic?

The genetic factors are completely laid down when the parental sperm and ovum unite. Thus the individual's genotype, by definition, is "fixed" at the moment of conception. Of course, different potentials of the genotype may be expressed at different times in the course of the individual's development. But beyond conception, whatever we observe or measure of the organism is a phenotype, and this, by definition, is *not* "fixed." The phenotype is a result of the organism's internal genetic mechanisms established at conception and all the physical and social influences that impinge on the organism throughout the course of its development. Intelligence is a phenotype, not a genotype, so the argument about whether or not intelligence is "fixed" is seen to be spurious.

The really interesting and important question, which can be empirically answered by the methods of quantitative genetics, is: what is the correlation between genotypes and phenotypes at any given point indevelopment? For continuous or metrical characteristics such as height and intelligence, the correlation, of course, can assume any value between 0 and 1. The square of the correlation between genotype and phenotype is technically known as the *heritability* of the characteristic, a concept which is discussed more fully in a later section.

The Stability of Intelligence Measures. The second aspect of the issue of "fixed intelligence" concerns the stability of intelligence measurements throughout the course of the individual's development. Since intelligence test scores are not points on an absolute scale of measurement like height and weight, but only indicate the individual's relative standing with reference to a normative population, the question we must ask is: To what extent do individuals maintain their standing relative to one another in measured intelligence over the course of time? The answer is to be found in the correlation between intelligence test scores on a group of persons at two points in time. Bloom (1964) has reviewed the major studies of this question and the evidence shows considerable consistency.

In surveying all the correlations reported in the literature between intelligence measured on the same individuals at two points in time, I have worked out a simple formula that gives a "best fit" to all these data. The formula has the virtue of a simple mnemonic, being much easier to remember than all the tables of correlations reported in the literature and yet being capable of reproducing the correlations with a fair degree of accuracy.

$$r_{12} = r_{tt} \text{ sqrt } (CA_1 / CA_2)$$

Equation 1.

where

ρ_{12}	=	the estimated correlation between tests given at times 1 and 2.
r_{tt}	=	the equivalent forms or immediate test-retest reliability of the test.

CA_1 = the subject's chronological age at the time of the first test.

CA_2 = the subject's chronological age at the time of the second test.

Limitation: The formula holds only up to the point where CA_2 is age 10, at which time the empirical value of r_{12} approaches an asymptote, showing no appreciable increase thereafter. Beyond age 10, regardless of the interval between tests, the obtained test-retest correlations fall in the range between the test's reliability and the square of the reliability (i.e., $r_{tt} > r_{12} > r_{tt}^2$). These simple generalizations are intended simply as a means of summarizing the mass of empirical findings. The accord with Bloom's conclusion, based on his thorough survey of the published evidence, that beyond age 8, correlations between repeated tests of general intelligence, corrected for the unreliability of measurement, are between +.90 and unity (Bloom, 1964, p. 61).

What these findings mean is that the IQ is not constant, but, like all other developmental characteristics, is quite variable early in life and becomes increasingly stable throughout childhood. By age 4 or 5, the IQ correlates about .70 with IQ at age 17, which means that approximately half (i.e., the square of the correlation) of the variance in adult intelligence can be predicted as early as age 4 or 5. This fact that half the variance in adult intelligence can be accounted for by age 4 has led to the amazing and widespread, but unwarranted and fallacious, conclusion that persons develop 50 percent of their mature intelligence by age 4! This conclusion, of course, does not at all logically follow from just knowing the magnitude of the correlation. The correlation between *height* at age 4 and at age 17 is also about .70, but who would claim that the square of the correlation indicated the proportion of adult height attained by age 4? The absurdity of this non sequitur is displayed in the prediction it yields: the average 4 year old boy should grow up to be 6 ft. 7 in. tall by age 17!

Intelligence has about the same degree of stability as other developmental characteristics. For example, up to age 5 or 6, height is somewhat more stable than intelligence, and thereafter the developmental rates of height and intelligence are about equally stable, except for a period of 3 or 4 years immediately after the onset of puberty, during which height is markedly less stable than intelligence. Intelligence is somewhat more stable than total body weight over the age range from 2 to 18 years. Intelligence has a considerably more stable growth rate than measures of physical strength (Bloom, 1964, pp. 46-47). Thus, although the IQ is certainly not "constant," it seems safe to say that under normal environmental conditions it is at least as stable as developmental characteristics of a strictly physical nature.

Intelligence as a Component of Mental Ability

The term "intelligence" should be reserved for the rather specific meaning I have assigned to it, namely, the general factor common to standard tests of intelligence. Any one verbal definition of this factor is really inadequate, but, if we must define it in so many words, it is probably best thought of as a capacity for abstract reasoning and problem solving.

What I want to emphasize most, however, is that *intelligence* should not be regarded as completely synonymous with what I shall call *mental ability,* a term which refers to the totality of a person's mental capabilities. Psychologists know full well that what they mean by intelligence in the technical sense is only a part of the whole spectrum of human abilities. The notion that a person's intelligence, or some test measurement thereof, reflects the totality of all that he can possibly do with his "brains" has long caused much misunderstanding and needless dispute. As I have already indicated, the particular constellations of abilities we now call "intelligence," and which we can measure by means of "intelligence" tests, has been singled out from the total galaxy of mental abilities as being especially important in our society mainly because of the nature of our traditional system of formal education and the occupational structure with which it is coordinated. Thus, the predominant importance of intelligence is derived, not from any absolute criteria or God-given desiderata, but from social demands. But neither does this mean, as some persons would like to believe, that intelligence exists only "by definition" or is merely an insubstantial figment of psychological theory and test construction. Intelligence fully meets the usual scientific criteria for being regarded as an aspect of objective reality, just as much as do atoms, genes, and electromagnetic fields. Intelligence has indeed been singled out as especially important by the educational and occupational demands prevailing in all industrial societies, but it is nevertheless a biological reality and not just a figment of social convention. Where educators and society in general are most apt to go wrong is in failing fully to recognize and fully to utilize a broader spectrum of abilities than just that portion which psychologists have technically designated as "intelligence." But keep in mind that it is this technical meaning of "intelligence" to which the term specifically refers throughout the present article.

The Distribution of Intelligence

Intelligence tests yield numerical scores or IQs (intelligence quotients) which are assumed to be, and in fact nearly are, "normally" distributed in the population. That is, the distribution of IQs conforms to the normal or so-called Gaussian distribution, the familiar "bell-shaped curve." The IQ, which is now the most universal "unit" in the measurement of intelligence, was originally defined as the ratio of the individual's mental age (MA) to his chronological age (CA):

$$IQ = (MA/CA) \times 100$$

(Beyond about 16 years of age, the formula ceases to make sense.) Mental age was simply defined as the typical or average score obtained on a test by children of a given age, and thus the average child by definition has an IQ of 100. Because of certain difficulties with the mental age concept, which we need not go into here, modern test constructors no longer attempt to measure mental age but instead convert raw scores (i.e., the number of test items gotten "right") directly into IQs for each chronological age group. The average IQ at each age is arbitrarily set at 100, and the IQ is defined as a normally distributed variable with a mean of 100 and a standard deviation of 15 points. (The standard deviation is an index of the amount of dispersion of scores; in the normal distribution 99.7 percent of the scores fall within +/- 3 standard deviations [i.e., +/- 45 IQ points] of the mean.)

There is really nothing mysterious about the fact that IQs are "normally" distributed, but it is not quite sufficient, either, to say that the normality of the distribution is just an artifact of test construction. There is a bit more to it than that.

Toss a hundred or so pennies into the air and record the number of heads that come "up" when they fall. Do this several thousand times and plot a frequency distribution of the number of heads that come up on each of the thousands of throws. You will have a distribution that very closely approximates the normal curve, and the more times you toss the hundred pennies the closer you will approximate the normal distribution.

Now, a psychological test made up of 100 or so items would behave in the same manner as the pennies, and produce a perfectly normal distribution of scores, if (*a*) the items have an average difficulty level of 1/2 [i.e., exactly half of the number of persons taking the test would get the item "right"], and (*b*) the items are independent, that is, all the interitem correlations are zero. Needless to say, no psychological test that has ever been constructed meets these "ideal" criteria, and this is just as well, for if we succeeded in devising such a test it would "measure" absolutely nothing but chance variation. If the test is intended to measure some trait, such as general intelligence, it will be impossible for all the test items to be completely uncorrelated. They will necessarily have some degree of positive correlation among them. Then, if the items are correlated, and if we still want the test to spread people out over a considerable range of scores, we can achieve this only if the items vary in level of difficulty; they cannot all have a difficulty level of 1/2. (Imagine the extreme case in which all item intercorrelations were perfect and the difficulty level of all items was 1/2. Then the "distribution" of scores would have only two points: half the testees would obtain a score of zero and half would obtain a perfect score.) So we need to have test items which have an *average* difficulty level of 1/2 in the test overall, but which cover a considerable range of difficulty levels, say, from .1 to .9. Thus, test constructors make up their tests of items which have rather low average intercorrelations (usually between .1 and .2) and a considerable range of difficulty levels. These two sets of conditions working together, then, yield a distribution of test scores in the population which is very close to "normal." So far it appears as though we have simply made our tests in such a way as to *force* the scores to assume a normal distribution. And that is exactly true.

But the important question still remains to be answered: is intelligence itself—not just our measurements of it—really normally distributed? In this form the question is operationally meaningless, since, in order to find the form of the distribution of intelligence, we first have to measure it, and we have constructed our measuring instruments in such a way as to yield a normal distribution. The argument about the distribution of intelligence thus appears to be circular. Is there any way out? The only way I know of is to look for evidence that our intelligence scales or IQs behave like an "interval scale." On an interval scale, the interval between any two points is equal to the interval between any other two points the same distance apart. Thus, intervals on the scale are equal and additive. If we *assume* that intelligence is "really" normally distributed in the population, and then measure it in such a way that we obtain a normal distribution of scores, our measurements (IQs) can be regarded as constituting an interval scale. If, then, the scale in fact behaves like an interval scale, there is some justification for saying that intelligence itself (not just IQ) is normally distributed. What evidence is there of the IQ's behaving like an interval scale? The most compelling evidence, I believe, comes from studies of the inheritance of intelligence, in which we examine the pattern of intercorrelations among relatives of varying degrees of kinship.

But, first, to understand what is meant by "behaving" like an interval scale, let us look at two well-known interval scales, the Fahrenheit and centigrade thermometers. We can prove that these are true intervale scales by showing that they "behave" like interval scales in the following manner: Mix a pint of ice water at 0°C with a pint of boiling water at 100°C. The resultant temperature of the mixture will be 50°C. Mix 3 pints of ice water with 1 pint of boiling water and the temperature of the mix will be 25°C. And we can continue in this way, mixing various proportions of water at different temperatures and predicting the resultant temperatures on the assumption of an interval scale. To the extent that the thermometer readings fit the predictions, they can be considered an interval scale.

Physical stature (height) is measured on an interval scale (more than that, it is also a ratio scale) in units which are independent of height, so the normal distribution of height in the population is clearly a fact of nature and not an artifact of the scale of measurement. A rather simple genetic model "explains" the distribution of height by hypothesizing that individual variations in height are the result of a large number of independent factors each having a small effect in determining stature. (Recall the penny-tossing analogy.) This model predicts quite precisely the amount of "regression to the population mean" of the children's average height from the parent's average height, a phenomenon first noted by Sir Francis Galton in 1885. The amount of "regression to the mean" from grandparent to grandchild is exactly double that from parent to child. These regression lines for various degrees of kinship are perfectly rectilinear throughout the entire range, except at the very lower end of the scale of height, where one finds midgets and dwarfs. The slope of the regression line changes in discrete jumps according to the remoteness of kinship of the groups being compared. All this could happen only if height were measured on an interval scale. The regression lines would not be rectilinear if the trait (height) were not measured in equal intervals.

Now, it is interesting that intelligence measurements show about the same degree of "filial regression," as Galton called it, that we find for height. The simple polygenic model for the inheritance of height fits the kinship correlations obtained for intelligence almost as precisely as it does for height. And the kinship regression lines are as rectilinear for intelligence as for height, throughout the IQ scale, except at the very lower end, where we find pathological types of mental deficiency analogous to midgets and dwarfs on the scale of physical stature. In brief, IQs behave just about as much like an interval scale as do measurements of height, which we know for sure is an interval scale. Therefore, it is not unreasonable to treat the IQ as an interval scale.

Although standardized tests such as the Stanford-Binet and the Wechsler Scales were each constructed by somewhat different approaches to achieving interval scales, they both agree in revealing certain systematic discrepancies from a perfectly normal distribution of IQs when the tests are administered to a very large and truly random sample of the population. These slight deviations of the distribution of IQs from perfect normality have shown up in many studies using a variety of tests. The most thorough studies and sophisticated discussions of their significance can be found in articles by Sir Cyril Burt (1957, 1963). The evidence, in short, indicates that intelligence is *not* distributed quite normally in the population. The distribution of IQs approximates normality quite closely in the IQ range from about 70 to 130. But outside this range there are slight, although very significant, departures from normality. From a scientific standpoint, these discrepancies are of considerable interest as genuine phenomena needing explanation.

Figure 1 shows an idealized distribution of IQs if they were distributed perfectly normally. Between IQ 70 and 130, the percentage of cases falling between different IQ intervals, as indicated in Figure 1, are very close to the actual percentages estimated from large samples of the population and the departures are hardly enough to matter from any practical standpoint.

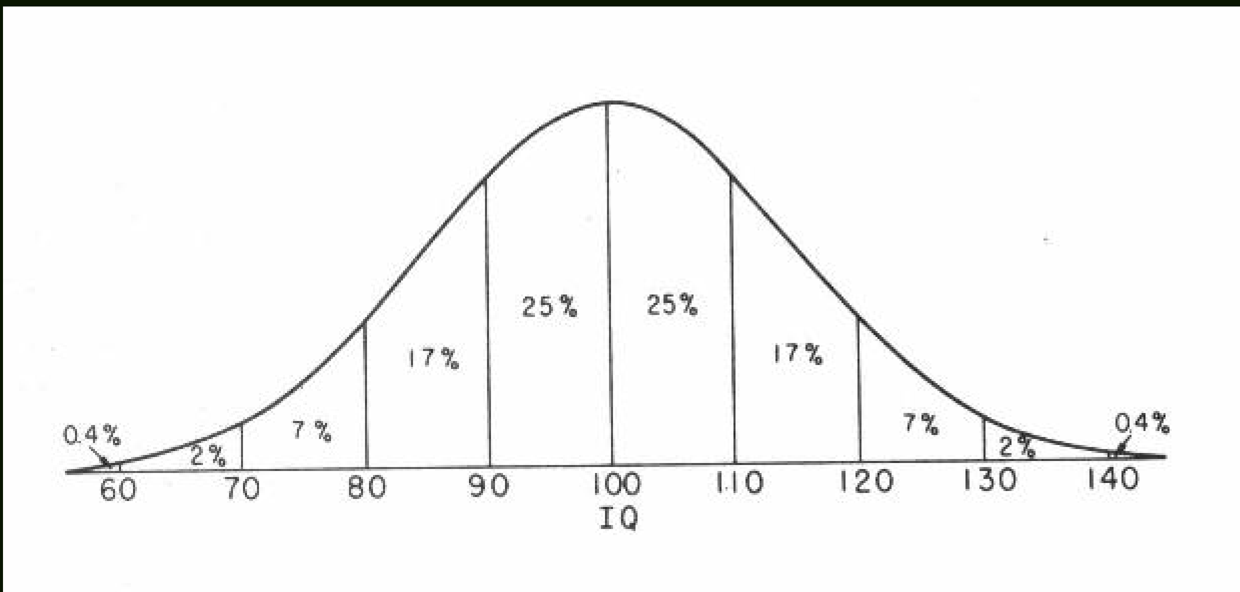

FIGURE 1.
The theoretical normal or Gaussian distribution of IQs, showing the expected percentages of the population in each IQ range. Except at the extremes (below 70 and above 130) these percentages are very close to actual population values. (The percentage figures total slightly more than 100% because of rounding.)

Examination of this normal curve can be instructive if one notes the consequences of shifting the total distribution up or down the IQ scale. The consequences of a give shift become more extreme out toward the "tails" of the distribution. For example, shifting the mean of the distribution from 100 down to 90 would put 50 percent instead of only 25 percent of the population below IQ 90; and it would put 9 percent instead of 2 percent below IQ 70. And in the upper tail of the distribution, of course, the consequences would be the reverse; instead of 25 percent above IQ 110, there would be only 9 percent, and so on. The point is that relatively small shifts in the mean of the IQ distribution can result in very large differences in the proportions of the population that fall into the very low or the very high ranges of intelligence. A 10 point downward shift in the mean, for example, would more than triple the percentage of mentally retarded (IQs below 70) in the population and would reduce the percentage of the intellectually "gifted" (IQs above 130) to less than one-sixth of their present number. It is in these tails of the normal distribution that differences become most conspicuous between various groups in the population that show mean IQ differences, for whatever reason, of only a few IQ points. From a knowledge of relatively slight mean differences between various social class and ethnic groups, for example, one can estimate quite closely the relatively large differences in their proportions in special classes for the educationally retarded and for the "gifted" and in the percentages of different groups receiving scholastic honors at graduation. It is simply a property of the normal distribution that the effects of group differences in the mean are greatly magnified in the different proportions of each group that we find as we move further out toward the upper or lower extremes of the distribution.

I indicated previously that the distribution of intelligence is really not quite "normal," but show certain systematic departures from "normality." These departures from the normal distribution are shown in Figure 2 in a slightly exaggerated form to make them clear. The shaded area is the normal distribution; the heavy line indicates the actual distribution of IQs in the population. We note that there are more very low IQs than would be expected in a truly normal distribution and also there is an excess of IQs at the upper end of the scale. Note, too, the slight excess in the IQ range between about 70 and 90.

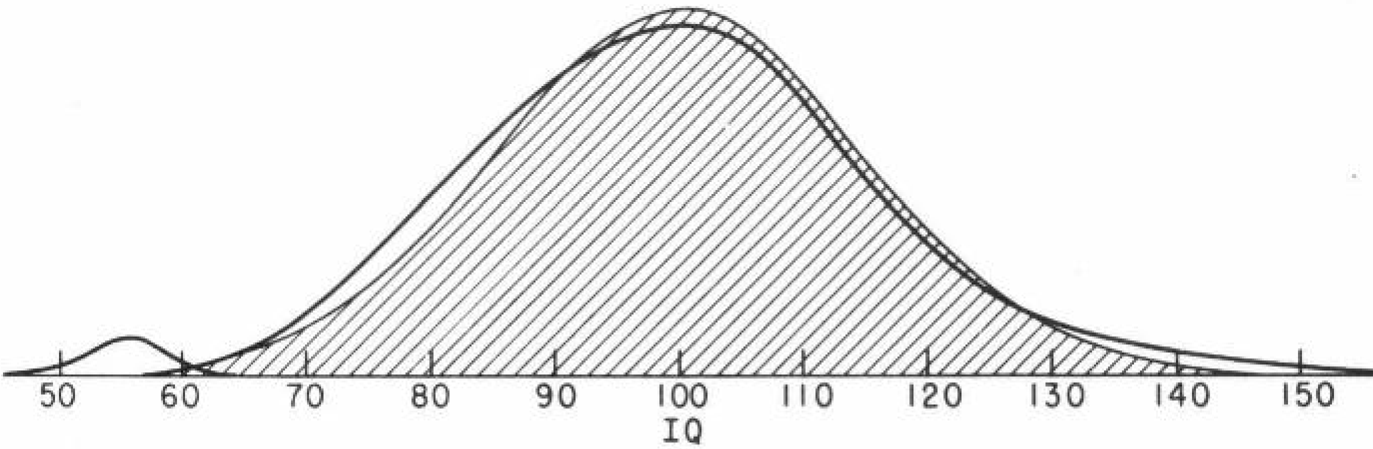

FIGURE 2.
Theoretical "normal" distribution of IQs (shaded curve) and the actual distribution in the population (heavy line), with the lower hump exaggerated for explanatory purposes. See text for explanation.

The very lowest IQs, below 55 or 60, we now know, really represent a different distribution from that of the rest of the intelligence distribution (Roberts, 1952; Zigler, 1967). Whatever factors are responsible for individual differences in the IQ range above 60 are not sufficient to account for IQs below this level, and especially below IQ 50. Practically all IQs below this level represent severe mental deficiency due to pathological conditions, massive brain damage, or rare genetic and chromosomal abnormalities. Only about 1/2 to 3/4 of 1 percent of the total population falls into the IQ range below 50; this is fewer than 1/3 of all individuals classed as mentally retarded (IQs below 70). These severe grades of mental defect are not just the lower extreme of normal variation. Often they are due to a single recessive or mutant gene whose effects completely override all the other genetic factors involved in intelligence; thus they have been called "major gene" defects. In this respect, the distribution of intelligence is directly analogous to the distribution of stature. Short persons are no more abnormal than are average or tall persons; all are instances of normal variation. But extremely short persons at the very lower end of the distribution are really part of another, abnormal, distribution, generally consisting of midgets and dwarfs. They are clearly not a part of normal variation. One of the commonest types of dwarfism, for example, is known to be caused by a single recessive gene.

Persons with low IQs caused by major gene defects or chromosomal abnormalities, like mongolism, are also usually abnormal in physical appearance. Persons with moderately low IQs that represent a part of normal variation, the so-called "familial mentally retarded," on the other hand, are physically indistinguishable from persons in the higher ranges if IQ. But probably the strongest evidence we have that IQs below 50 are a group apart from the mildly retarded, who represent the lower end of normal variation, comes from comparisons of the siblings of the severely retarded with siblings of the mildly retarded. In England, where this has been studied intensively, these two retardate groups are called imbecile (IQs below 50) and feebleminded (IQs 50 to 75). Figure 3 shows the IQ distributions of the *siblings* of imbecile and feebleminded children (Roberts, 1952). Note that the siblings of imbeciles have a much higher average level of intelligence than the siblings of the feebleminded. The latter group, furthermore, shows a distribution of IQs that would be predicted from a genetic model intended to account for the normal variation of IQ in the population. This model does not at all predict the IQ distribution for the imbecile sibships. To explain the results shown in Figure 3 one must postulate some additional factors (gene or chromosome defects, pathological conditions, etc.) that cause imbecile and idiot grades of mental deficiency.

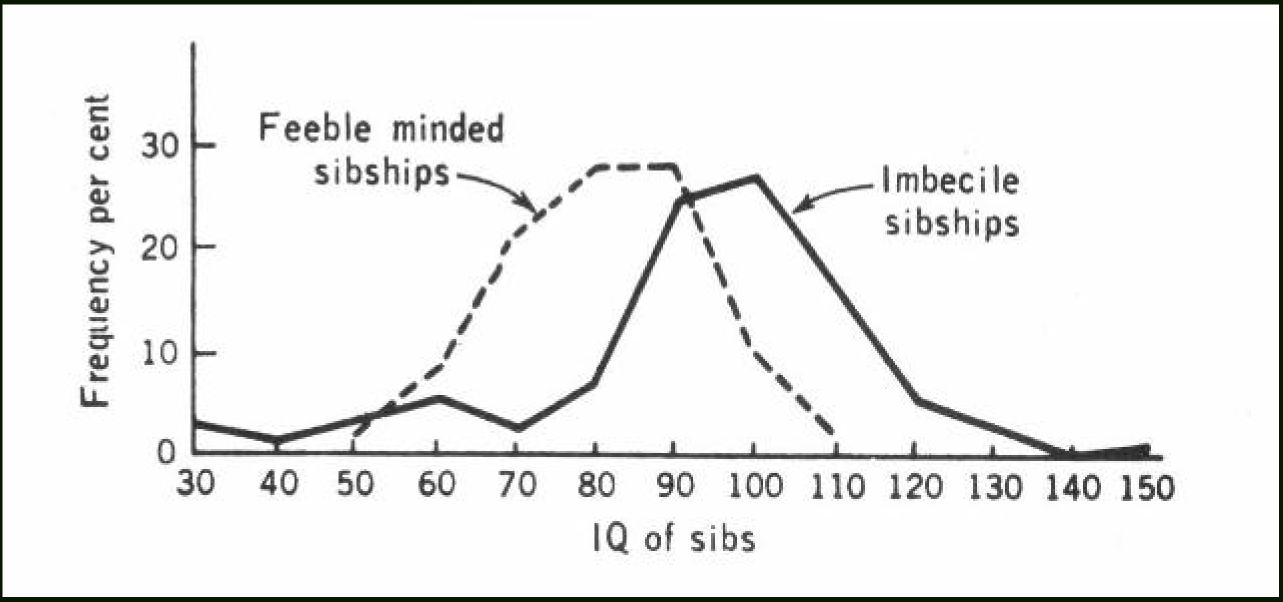

FIGURE 3.
Frequency distributions of the IQs of sibs of feebleminded and imbeciles of the IQ range 30-68. (Roberts, 1952.)

Another interesting point of contrast between severe mental deficiency and mild retardation is the fact noted by Kushlick (1966, p. 130), in surveying numerous studies, that "The parents of severely subnormal children are evenly distributed among all the social strata of industrial society, while those of mildly subnormal subjects come predominantly from the lower social classes. There is now evidence which suggests that mild subnormality in the absence of abnormal neurological signs (epilepsy, electroencephalographic abnormalities, biochemical abnormalities, chromosomal abnormalities or sensory defects) is virtually confined to the lower social classes. Indeed, there is evidence that almost no children of higher social class parents have IQ scores less than 80, unless they have one of the pathological processes mentioned above."

In the remainder of this article we shall not be further concerned with these exceptionally low IQs below 50 or 60, which largely constitute a distribution of abnormal conditions superimposed on the factors that make for normal variation in intelligence. We shall be mainly concerned with the factors involved in the normal distribution.

Returning to Figure 2, the best explanation we have for the "bulge" between 70 and 90 is the combined effects of severe environmental disadvantages and of emotional disturbances that depress test scores. Burt (1963) has found that when, independent of the subjects' test performance there is evidence for the existence of factors that depress performance, and these exceptional subjects' scores are removed from the distribution, this "bulge" in the 70-90 range is diminished or erased. Also, on retest under more favorable conditions, the IQs of many of these exceptional subjects are redistributed at various higher points on the scale, thereby making the IQ distribution more normal.

The "excess" of IQs at the high end of the scale is certainly a substantial phenomenon, but it has not yet been adequately accounted for. In his multifactorial theory of the inheritance of intelligence, Burt (1958) has postulated major gene effects that make for exceptional intellectual abilities represented at the upper end of the scale, just as other major gene effects make for the subnormality found at the extreme lower end of the scale. One might also hypothesize that superior genotypes for intellectual development are pushed to still greater superiority in their phenotypic expression through interaction with the environment. Every recognition of superiority leads to its greater cultivation and encouragement by the individual's social environment. This influence is keenly evident in the developmental histories of persons who have achieved exceptional eminence (Goertzel & Goertzel, 1962). Still another possible explanation of the upper-end "excess" lies in the effects of assortative mating in the population, meaning the tendency for "like to marry like." If the degree of resemblance in intelligence between parents in the upper half of the IQ distribution were significantly greater than the degree of resemblance of parents in the below average range, genetic theory would predict the relative elongation of the upper tail of the distribution. This explanation, however, must remain speculative until we have more definite evidence of whether there is differential assortative mating in different regions of the IQ distribution.

The Concept of Variance. Before going on to discuss the factors that account for normal variation in intelligence among individuals in the population, a word of explanation is in order concerning the quantification of variation. The amount of dispersion of scores depicted by the distributions in Figures 1 and 2 is technically expressed as the *variance,* which is the square of the standard deviation of the scores in the distribution. (Since the standard deviation of IQs in the population is 15, the total variance is 225.) *Variance* is a basic concept in all discussion of individual differences and population genetics. If you take the difference between every score and the mean of the total distribution, square each of these differences, sum them up, and divide the sum by the total number of scores, you have a quantity called the *variance.* It is an index of the total amount of variation among scores. Since variance represents variation on an additive scale, the total variance of a distribution of scores can be partitioned into a number of components, each one due to some factor which contributes a certain specifiable proportion of the variance, and all these variance components add up to the total variance. The mathematical technique for doing this, called "the analysis of variance," was invented by Sir Ronald Fisher, the British geneticist and statistician. It is one of the great achievements in the development of statistical methodology.

The Inheritance of Intelligence

"In the actual race of life, which is not to get ahead, but to get ahead of somebody, the chief determining factor is heredity." So said Edward L. Thorndike in 1905. Since then, the preponderance of evidence has proved him right, certainly as concerns those aspects of life in which intelligence plays an important part.

But one would get a quite different impression from reading most of the recent popular textbooks of psychology and education. Genetic factors in individual differences have usually been belittled, obscured, or denigrated, probably for reasons of interest mainly on historical, political, and ideological grounds which we need not go into here. Some of the following quotations, each from different widely used texts in our field, give some indication of the basis for my complaint. "We can attribute no particular portion of intelligence to heredity and no particular portion to the environment." "The relative influence of heredity and environment upon intelligence has been the topic of considerable investigations over the last half century. Actually the problem is incapable of solution since studies do not touch upon the problem of heredity and environment but simply upon the susceptibility of the content of a particular test to environmental influences." "Among people considered normal, the range of genetic variations is not very great." "Although at the present time practically all responsible workers in the field recognize that conclusive proof of the heritability of mental ability (where no organic or metabolic pathology is involved) is still lacking, the assumption that subnormality has a genetic basis continues to crop up in scientific studies." "There is no evidence that nature is more important than nurture. These two forces always operate together to determine the course of intellectual development." The import of such statements apparently filters up to high levels of policy-making, for we find a Commissioner of the U.S. Office of Education stating in a published speech that children "...all have similar potential at birth. The differences occur shortly thereafter." These quotations typify much of the current attitude toward heredity and environment that has prevailed in education in recent years. The belief in the almost infinite plasticity of intellect, the ostrich-like denial of biological factors in individual differences, and the slighting of the role of genetics in the study of intelligence can only hinder investigation and understanding of the conditions, processes, and limits through which the social environment influences human behavior.

But fortunately we are beginning to see some definite signs that this mistreatment of the genetic basis of intelligence by social scientists may be on the wane, and that a biosocial view of intellectual development more in accord with the evidence is gaining greater recognition. As Yale psychologist Edward Zigler (1968) has so well stated:

> Not only do I insist that we take the biological integrity of the organism seriously, but it is also my considered opinion that our nation has more to fear from unbridled environmentalists than they do from those who point to such integrity as one factor in the determination of development. It is the environmentalists who have been writing review after review in which genetics are ignored and the concept of capacity is treated as a dirty word. It is the environmentalists who have placed on the defensive any thinker who, perhaps impressed by the revolution in biological thought stemming from the discoveries involving RNA-DNA phenomena, has had the temerity to suggest that certain behaviors may be in part the product of the read-out mechanism residing within the programmed organism. It is the unbridled environmentalist who emphasizes the plasticity of the intellect, that tells us one can change both the general rate of development and the configuration of intellectual processes which can be referred to as the intellect, if we could only subject human beings to the proper technologies. In the educational realm, this has spelled itself out in the use of panaceas, gadgets, and gimmicks of the most questionable sort. It is the environmentalist who suggests to parents how easy it is to try to raise the child's IQ and who has prematurely led many to believe that the retarded could be made normal, and the normal made geniuses. It is the environmentalist who has argued for pressure-cooker schools, at what psychological cost, we do not yet know.

Most geneticists and students of human evolution have fully recognized the role of culture in shaping "human nature," but also they do not minimize the biological basis of diversity in human behavioral characteristics. Geneticist Theodosius Dobzhansky (1968, p. 554) has expressed this viewpoint in the broadest terms: "The trend of cultural evolution has been not toward making everybody have identical occupations but toward a more and more differentiated occupational structure. What would be the most adaptive response to this trend? Certainly nothing that would encourage genetic uniformity.... To argue that only environmental circumstances and training determine a person's behavior makes a travesty of democratic notions of individual choice, responsibility, and freedom."

Evidence from Studies of Selective Breeding

The many studies of selective breeding in various species of mammals provides conclusive evidence that many behavioral characteristics, just as most physical characteristics, can be manipulated by genetic selection (see Fuller & Thompson, 1962; Scott and Fuller, 1965). Rats, for example, have been bred for maze learning ability in many different laboratories. It makes little difference whether one refers to this ability as rat "intelligence," "learning ability" or some other term—we know that it is possible to breed selectively for whatever the factors are that make for speed of maze learning. To be sure, individual variation in this complex ability may be due to any combination of a number of characteristics involving sensory acuity, drive level, emotional stability, strength of innate turning preferences, brain chemistry, brain size, structure of neural connections, speed of synaptic transmission, or whatever. The point is that the molar behavior of learning to get through a maze efficiently without making errors (i.e., going up blind alleys) can be markedly influenced in later generations by selective breeding of the parent generations of rats who are either fast or slow ("maze bright" or "maze dull," to use the prevailing terminology in this research) in learning to get through the maze. Figure 4 shows the results of one such genetic selection experiment. They are quite typical; with only six generations of selection the offspring of the "dull" strain make 100 percent more errors in learning the maze than do the offspring of the "bright" strain (Thompson, 1954). In most experiments of this type, of course, the behaviors that respond so dramatically to selection are relatively simple as compared with human intelligence, and the experimental selection pressure is severe, so the implications of such findings for the study of human variation should not be overdrawn. Yet geneticists seem to express little doubt that many behavioral traits in humans would respond similarly to genetic selection. Three eminent geneticists (James F. Crow, James V. Neel, and Curt Stern) of the National Academy of Sciences recently prepared a "position statement," which was generally hedged by extreme caution and understatement, that asserted: "Animal experiments have shown that almost any trait can be changed by selection. ... A selection program to increase human intelligence (or whatever is measured by various kinds of 'intelligence' tests) would almost certainly be successful in some measure. The same is probably true for other behavioral traits. The *rate* of increase would be somewhat unpredictable, but there is little doubt that there would be progress" (National Academy of Sciences, 1967, p. 893).

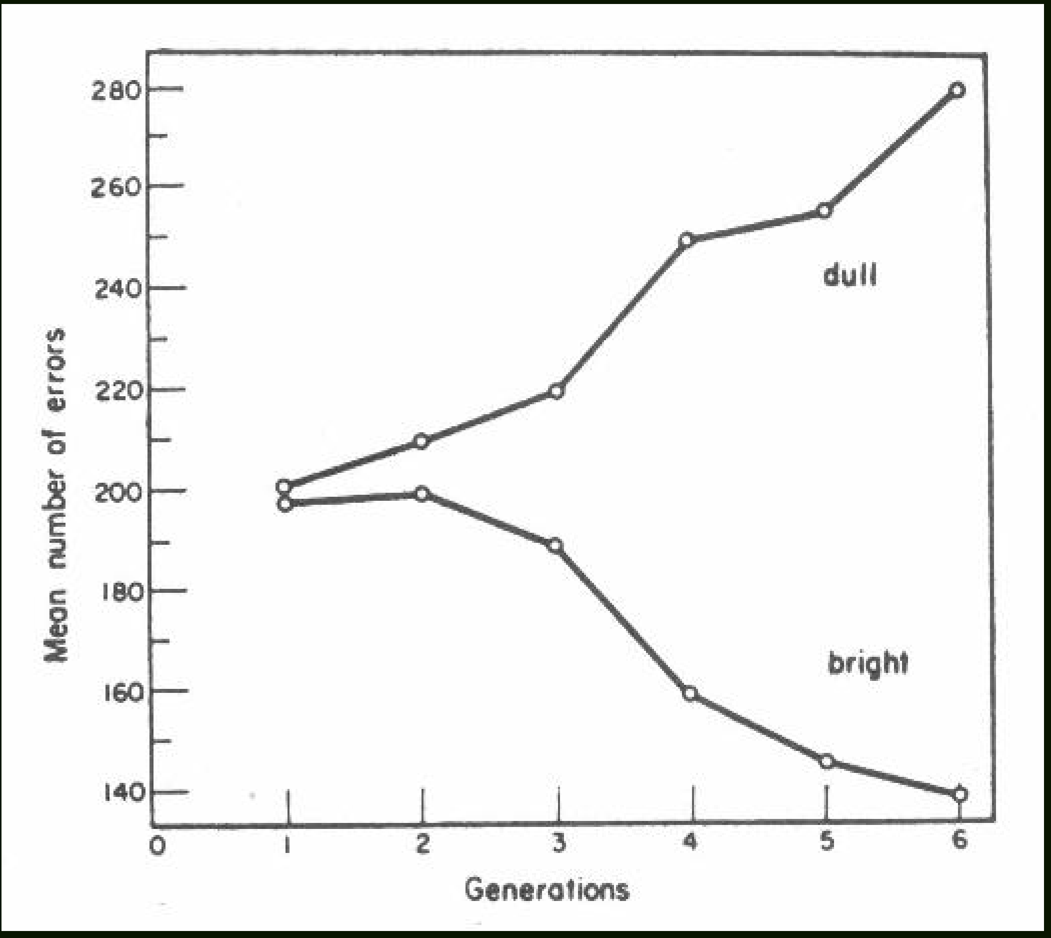

FIGURE 4.
The mean error scores in maze learning for successive generations of selectively bred "bright" and "dull" strains of McGill rats. (After Thompson, 1954.)

Direct Evidence of Genetic Influences on Human Abilities

One of the most striking pieces of evidence for the genetic control of mental abilities is a chromosomal anomaly called Turner's syndrome. Normal persons have 46 chromosomes. Persons with Turner's syndrome have only 45. When their chromosomes are stained and viewed under the microscope, it is seen that the sex-chromatin is missing from one of the two chromosomes that determine the individual's sex. In normal persons this pair of chromosomes is conventionally designated by XY for males and XX for females. The anomaly of Turner's syndrome is characterized by XO. These persons always have the morphologic appearance of females but are always sterile, and they show certain physical characteristics such as diminutive stature, averaging about five feet tall as adults. The interesting point about Turner's cases from our standpoint is that although their IQs on most verbal tests of intelligence show a perfectly normal distribution, their performance on tests involving spatial ability or perceptual organization is abnormally low (Money, 1964). Their peculiar deficiency in spatial perceptual ability is sometimes so severe as to be popularly characterized as "space-form blindness." It is also interesting that Turner's cases seem to be more or less uniformly low on spatial ability regardless of their level of performance on other tests of mental ability. These rare persons also report unusual difficulty with arithmetic and mathematics in school despite otherwise normal or superior intelligence. So here is a genetic aberration, clearly identifiable under the microscope, which has quite specific consequences on cognitive processes. Such specific intellectual deficiencies are thus entirely possible without there being any specific environmental deprivations needed to account for them.

There are probably other more subtle cognitive effects associated with the sex chromosomes in normal persons. It has long been suspected that males have greater environmental vulnerability than females, and Nancy Bayley's important longitudinal research on children's mental development clearly shows both a higher degree and a greater variety of environmental and personality correlates of mental abilities in boys than in girls (Bayley, 1965b, 1966, 1968).

Polygenic Inheritance

Since intelligence is basically dependent on the structural and biochemical properties of the brain, it should not be surprising that differences in intellectual capacity are partly the result of genetic factors which conform to the same principles involved in the inheritance of physical characteristics. The general model that geneticists have devised to account for the facts of inheritance of continuous or metrical physical traits, such as stature, cephalic index, and fingerprint ridges, also applies to intelligence. *The mechanism of inheritance for such traits is called polygenic, since normal variation in the characteristic is the result of multiple genes whose effects are small, similar, and cumulative.* The genes can be thought of as the pennies in the coin-tossing analogy described previously. Some genes add a positive increment to the metric value of the characteristic ("heads") and some genes add nothing ("tails"). The random segregation of the parental genes in the process of gametogenesis (formation of the sex cells) and their chance combination in the zygote (fertilized egg) may be likened to the tossing of a large number of pennies, with each "head" adding a positive increment to the trait, thereby producing the normal bell-shaped distribution of trait values in a large number of tosses. The actual number of genes involved in intelligence is not known. In fact, the total number of genes in the human chromosomes is unknown. The simplest possible model would require between ten and twenty gene pairs (alleles) to account for the normal distribution of intelligence, but many more genes than this are most likely involved (Gottesman, 1963, pp. 290-291).

The Concept of Heritability

The study of the genetic basis of individual differences in intelligence in humans has evolved in the traditions and methods of that branch of genetics called quantitative genetics or population genetics, the foundations of which were laid down by British geneticists and statisticians such as Galton, Pearson, Fisher, Haldane, and Mather, and, in the United States, by J. L. Lush and Sewall Wright. Probably the most distinguished exponent of the application of these methods to the study of intelligence is Sir Cyril Burt, whose major writings on this subject are a "must" for students of individual differences (Burt, 1955, 1958, 1959, 1961, 1966; Burt & Howard, 1956, 1957).

One aim of this approach to the study of individual differences in intelligence is to account for the total variance in the population (excluding pathological cases at the bottom of the distribution) in terms of the proportions of the variance attributable to various genetic and environmental components. It will pay to be quite explicit about just what this actually means.

Individual differences in such measurements of intelligence as the IQ are represented as population variance in a phenotype $V_{I'}$, and are distributed approximately as shown in Figure 1. Conceptually, this total variance of the phenotypes can be partitioned into a number of variance components, each of which represents a source of variance. The components, of course, all add up to the total variance. Thus,

$$V_P = \{ (V_G + V_{AM}) + V_D + V_i \} / V_H + \{ V_E + 2COV_{HE} + V_I \} / V_E + V_e$$

Equation 2.

where

V_P	=	phenotypic variance in the population
V_G	=	genic (or additive) variance
V_{AM}	=	variance due to assortative mating. V_{AM}=0 under random mating (panmixia).
V_D	=	dominance deviation variance
V_i	=	epistatis (interaction among genes at 2 or more loci)
V_E	=	environmental variance
COV_{HE}	=	covariance of heredity and environment
V_I	=	true statistical interaction of genetic and environmental factors
V_e	=	error of measurement (unreliability)

Here are a few words of explanation about each of these variance components.

Phenotypic Variance. V_P is already clear; it is the total variance of the trait measurements in the population.

Genic Variance, V_G, the genic (or additive) variance, is attributable to gene effects which are additive; that is, each gene adds an equal increment to the metric value of the trait. Sir Ronald Fisher referred to this component a "the essential genotypes," since it is the part of the genetic inheritance which "breeds true"—it accounts for the resemblance between parents and offspring. If trait variance involved nothing but additive genic effects, the average value of all the offspring that could theoretically be born to a pair of parents would be exactly equal to the average value of the parents (called the midparent value). It is thus the genic aspect which is most important to agriculturists and breeders of livestock, since it is the genetic component of the phenotype variance that responds to selection according to the simple rule of "like begets like." The larger the proportion of genic variance involved in a given characteristic, the fewer is the number of generations of selective breeding required to effect a change of some specified magnitude in the characteristic.

Assortative Mating. V_{AM}, the variance due to assortative mating, is conventionally not separated from V_G, since assortative mating actually affects the properties of V_G directly. I have separated these components here for explanatory reasons, and it is, in fact, possible to obtain independent estimates of the two components. If mating were completely random in the population with respect to a given characteristic—that is, if the correlation between parents were zero (a state of affairs known as *panmixia*)—the V_{AM}

component would also be equal to zero and the population variance on the trait in question would therefore be reduced.

Assortative mating has the effect of increasing the resemblance among siblings and also of increasing the differences between families in the population. (In the terminology of analysis of variance, assortative mating decreases the *Within* Families variance and increases the *Between* Families variance.)

For some human characteristics the degree of assortative mating is effectively zero. This is true of fingerprint ridges, for example. Men and women are obviously not attracted to one another on the basis of their fingerprints. Height, however, has an assortative mating coefficient (i.e., the correlation between mates) of about .30. The IQ, interestingly enough, shows a higher degree of assortative mating in our society than any other measurable human characteristic. I have surveyed the literature on this point, based on studies in Europe and North America, and find that the correlation between spouses' intelligence test scores averages close to +.60. Thus, spouses are more alike in intelligence than brothers and sisters, who are correlated about .50.

As Eckland (1967) has pointed out, this high correlation between marriage partners does not come about solely because men and women are such excellent judges of one another's intelligence, but because mate selection is greatly aided by the highly visible selective processes of the educational system and the occupational hierarchy. Here is a striking instance of how educational and social factors can have far-reaching genetic consequences in the population. One would predict, for example, that in preliterate or preindustrial societies assortative mating with respect to intelligence would be markedly less than it is in modern industrial societies. The educational screening mechanisms and socio-economic stratification by which intelligence becomes more readily visible would not exist, and other traits of more visible importance in the society would take precedence over intelligence as a basis for assortative mating. Even in our own society, there may well be differential degrees of assortative mating in different segments of the population, probably related to their opportunities for educational and occupational selection. When any large and socially insulated group is not subject to the social and educational circumstances that lead to a high degree of assortative mating for intelligence, there should be important genetic consequences. One possible consequence is some reduction of the group's ability, not as individuals but as a group, to compete intellectually. Thus probably one of the most cogent arguments for society's promoting full equality of educational, occupational, and economic opportunity lies in the possible genetic consequences of these social institutions.

The reason is simply that assortative mating increases the genetic variance in the population. By itself this will not affect the mean of the trait in the population, but it will have a great effect on the proportion of the population falling in the upper and lower tails of the distribution. Under present conditions, with an assortative mating coefficient of about .60, the standard deviation of IQs is 15 points. If assortative mating for intelligence were reduced to zero, the standard deviation of IQs would fall to 12.9. The consequences of this reduction in the standard deviation would be most evident at the extremes of the intelligence distribution. For example, assuming a normal distribution if IQs and the present standard deviation of 15, the frequency (per million) of persons above IQ 130 is 22,750. Without assortative mating the frequency of IQs over 130 would fall to 9,900, or only 43.5 percent of the present frequency. For IQs above 145, the frequency (per million) is 1,350 and with no assortative mating would fall to 241, or 17.9 percent of the present frequency. And there are now approximately 20 times as many persons above an IQ of 160 as we would find if there were no assortative mating for intelligence.[3] Thus differences in assortative mating can have a profound effect on a people's intellectual resources, especially at the levels of intelligence required for complex problem solving, invention, and scientific and technological innovation.

But what is the effect of assortative mating on the lower tail of the distribution? On theoretical grounds we should also expect it to increase the proportion of low IQs in the population. It probably does this to some extent, but not as much as it increases the frequency of higher IQs, because there is a longer-term consequence of assortative mating which must also be considered. A number of studies have shown that in populations practicing a high degree of assortative mating, persons below IQ 75 are much less successful in finding marriage partners and, as a group, have relatively fewer offspring than do persons of higher intelligence (Bajema, 1963, 1966; Higgins, Reed & Reed, 1962). Since assortative mating increases variance, it in effect pushes more people into the below IQ 75 group, where they fail to reproduce, thereby resulting in a net selection for genes favoring high intelligence. Thus, in the long run, assortative mating may have a eugenic effect in improving the general level of intelligence in the population.

Dominance Deviation. V_D, the dominance deviation variance, is apparent when we observe a systematic discrepancy between the average value of the parents and the average value of their offspring on a given characteristic. Genes at some of the loci in the chromosome are recessive (r) and their effects are not manifested in the phenotype unless they are paired with another recessive at the same locus. If paired with a dominant gene (D), their effect is overridden or "dominated" by the dominant gene. Thus, in terms of increments which genes add to the metric value of the phenotype, if r=0 and D=1, then r+r=0, and D+D=2, but D+r will equal 2, since D dominates r. Because of the presence of some proportion of recessive genes in the genotypes for a particular trait, not all of the parents' phenotypic characteristics will show up in their offspring, and, of course, vice versa: not all of the offspring's characteristics will be seen in the parents. This makes for a less than perfect correlation between midparent and midchild values on the trait in question. V_D, the dominance variance, represents the component of variance in the population which is due to this average discrepancy between parents and offspring. The magnitude of V_D depends upon the proportions of dominant and recessive genes constituting the genotypes for the characteristic in the population.

Epistasis. V_i is the variance component attributable to epistasis, which means the interaction of the effects among genes at two or more loci. When genes "interact," their effects are not strictly additive; that is to say, their combined effect may be more or less than the sum of their separate effects. Like dominance, epistasis also accounts for some of the lack of resemblance between parents and their offspring. And it increases the population variances by a component designated as V_i.

Environmental Variance. "Environmental" really means all sources of variance not attributable to genetic effects or errors of measurement (i.e., test unreliability). In discussions of intelligence, the environment is often thought of only in terms of the social and cultural influences on the individual. While these are important, they are not the whole of "environment," which includes other more strictly biological influences, such as the prenatal environment and nutritional factors early in life. In most studies of the heritability of intelligence "environment" refers to all variance that is not accounted for by genetic factors $[(V_G + V_{AM}) + V_D + V_i]$ and measurement errors (V_e).

Covariance of Heredity and Environment. This term can also be expressed as

$$2\ r_{HE}\ \mathrm{sqrt}(V_H\ V_E)$$

where r_{HE} is the correlation between heredity and environment, V_H is the variance due to all genetic factors, and V_E is variance due to all environmental factors. In other words, if there is a positive correlation between genetic and environmental factors, the population variance is increased by a theoretically specifiable amount indicated by the covariance term in Equation 2.

Such covariance undoubtedly exists for intelligence in our society. Children with better than average genetic endowment for intelligence have a greater than chance likelihood of having parents of better than average intelligence who are capable of providing environmental advantages that foster intellectual development. Even among children within the same family, parents and teachers will often give special attention and opportunities to the child who displays exceptional abilities. A genotype for superior ability may cause the social environment to foster the ability, as when parents perceive unusual responsiveness to music in one of their children and therefore provide more opportunities for listening, music lessons, encouragement to practice, and so on. A bright child may also create a more intellectually stimulating environment for himself in terms of the kinds of activities that engage his interest and energy. And the social rewards that come to the individual who excels in some activity reinforce its further development.

Thus the covariance term for any given trait will be affected to a significant degree by the kinds of behavioral propensities the culture rewards or punishes, encourages or discourages. For traits viewed as desirable in our culture, such as intelligence, heredity and environmental factors will be positively correlated. But for some other traits which are generally viewed as socially undesirable, heredity and environmental influences may be negatively correlated. This means that the social environment tends to discourage certain behavioral propensities when they are out of line with the values of the culture. Then, instead of heredity and environment acting in the same direction, they work in opposite directions, with a consequent reduction in the population variance in the trait. Overt aggressive tendencies may be a good example of behavior involving a negative correlation between genotypic propensities and environmental counter-pressures. An example of negative heredity-environment correlation in the scholastic realm would be found in the case where a child with a poor genetic endowment for learning some skill which is demanded by societal norms, such as being able to read, causes the child's parents to lavish special tutorial attention on their child in an effort to bring his performance up to par.

In making overall estimates of the proportions of variance attributable to hereditary and environmental factors, there is some question as to whether the covariance component should be included on the side of heredity or environment. But there can be no "correct" answer to this question. To the degree that the individual's genetic propensities cause him to fashion his own environment, given the opportunity, the covariance (or some part of it) can be justifiably regarded as part of the total heritability of the trait. But if one wishes to estimate what the heritability of the trait would be under artificial conditions in which there is absolutely no freedom for variation in individuals' utilization of their environment, then the covariance term should be included on the side of environment. Since most estimates of the heritability of intelligence are intended to reflect the existing state of affairs, they usually include the covariance in the proportion of variance due to heredity.

Interaction of Heredity and Environment. The *interaction* of genetic and environmental factors (V_I) must be clearly distinguished from the *covariance* of heredity and environment. There is considerable confusion concerning the meaning of interaction in much of the literature on heredity and intelligence. It is claimed, for example, that nothing can be said about the relative importance of heredity and environment because intelligence is the result of the "interaction" of these influences and therefore their independent effects cannot be estimated. This is simply false. The proportion of the population variance due to genetic x environment interaction is conceptually and empirically separable from other variance components, and its independent contribution to the total variance can be known. Those who call themselves "interactionists," with the conviction that they have thereby either solved or risen above the whole issue of the relative contributions of heredity and environment to individual differences in intelligence, are apparently unaware that the preponderance of evidence indicates that the interaction variance, V_I is the smallest component of the total phenotypical variance of intelligence.

What *interaction* really means is that different genotypes respond in different ways to the same environmental factors. For example, genetically different individuals having the same initial weight and the same activity level may gain weight at quite different rates all under exactly the same increase in caloric intake. Their genetically different constitutions cause them to metabolize exactly the same intake quite differently. An example of genotype x environmental interaction in the behavioral realm is illustrated in Figure 5. Strains of rats selectively bred for "brightness" or "dullness" in maze learning show marked differences in maze performance according to the degree of sensory stimulation in the conditions under which they are reared. For the "bright" strain, the difference between being reared in a "restricted" or in a "normal" environment makes a great difference in maze performance. But for the "dull" strain, the big difference is between a "normal" and a "stimulating" environment. While the strains differ greatly when reared under "normal" conditions (presumably the conditions under which they were selectively bred for "dullness" and "brightness"), they do not differ in the least when reared in a "restricted" environment and only slightly in a "stimulating" environment. This is the meaning of the genetic x environment interaction. Criticisms of the analysis of variance model for the components of phenotypic variance (e.g., Equation 2), put forth first by Loevinger (1943) and then by Hunt (1961, p. 329), are based on the misconception that the model implies that all effects of heredity and environment are strictly additive and there is no "non-additive" or interaction term. The presence of V_I in Equation 2 explicitly shows that the heredity x environment interaction is included in the analysis of variance model, and the contribution of V_I to the total variance may be estimated independently of the purely additive effects of heredity and environment. The magnitude of V_I for any given characteristic in any specified population is a matter for empirical study, not philosophic debate. If V_I turns out to constitute a relatively small proportion of the total variance, as the evidence shows is the case for human intelligence, this is not a fault of the analysis of variance model. It is simply a fact. If the interaction variance actually exists in any significant amount, the model will reveal it.

Several studies, reviewed by Wiseman (1964, p. 55; 1966, p. 66), provide most of the information we have concerning what may be presumed to be an heredity x environment interaction with respect to human intelligence. The general finding is that children who are more than one standard deviation (SD) above the mean IQ show greater correlations with environmental factors than do children who are more than one SD below the mean. In other words, if the heritability of IQ were determined in these two groups separately, it would be higher in the low IQ groups. Also, when siblings within the same family are grouped into above and below IQ 100, the scholastic achievement of the above 100 group shows a markedly higher correlation with environmental factors than in the below 100 group. This indicates a true interaction between intelligence and environment in determining educational attainments.

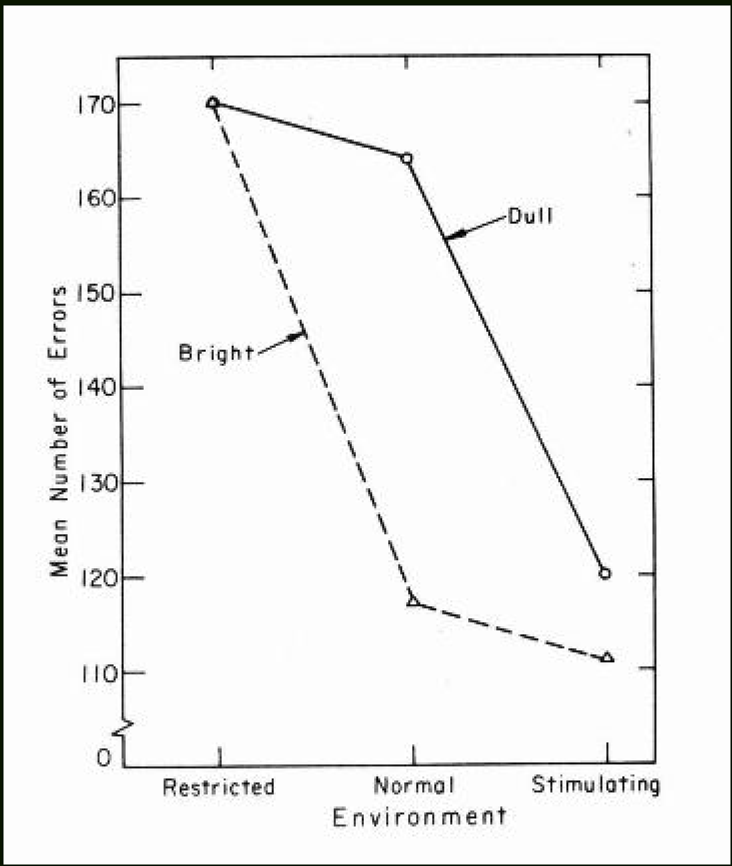

FIGURE 5.
Illustration of a true genotype x environment interaction for error scores in maze learning by "bright" and "dull" strains of rats raised in "restricted," "normal," and "stimulating" environments. (After Cooper & Zubek, 1958.)

Error Variance. The variance due to errors of measurement (V_e) is, of course, unwanted but unavoidable, since all measurements fall short of perfect reliability. The proportion of test score variance due to error is equal to $1-r_{tt}$ (where r_{tt} is the reliability of the test, that is, its correlation with itself). For most intelligence tests, error accounts for between 5 and 10 percent of the variance.

Definition of Heritability

Heritability is a technical term in genetics meaning specifically the proportion of phenotypic variance due to variance in genotypes. When psychologists speak of heritability they almost invariably define it as:

$$H = \{ (V_G + V_{AM}) + V_D + V_i \} / (V_P - V_e)$$

Equation 3.

Although this formula is technically the definition of H, heritability estimated in psychological studies may also include the covariance term of Equation 2 in the numerator of Equation 3.

Common Misconceptions About Heritability

Certain misconceptions about heritability have become so widespread and strongly ingrained that it is always necessary to counteract them before presenting the empirical findings on the subject, lest these findings only add to the confusion or provoke the dogmatic acceptance or rejection of notions that are not at all implied by the meaning of heritability.

Heredity versus Environment. Genetic and environmental factors are not properly viewed as being in opposition to each other. Nor are they an "all or none" affair. Any observable characteristic, physical or behavioral, is a phenotype, the very existence of which depends upon both genetic and environmental conditions. The legitimate question is not whether the characteristic is due to heredity or environment, but what proportion of the population variation in the characteristic is attributable to genotypic variation (which is H, the heritability) and what proportion is attributable to non-genetic or environmental variation in the population (which is 1-H). For metric characteristics like stature and intelligence, H can have values between 0 and 1.

Individual versus Population. Heritability is a population statistic, describing the relative magnitude of the genetic component (or set of genetic components) in the population variance of the characteristic in question. It has no sensible meaning with reference to a measurement of characteristic in an individual. A single measurement, by definition, has no variance. There is no way of partitioning a given individual's IQ into hereditary and environmental components, as if the person inherited, say, 80 points of IQ and acquired 20 additional points from his environment. This is, of course, nonsense. *The square root of the heritability, sqrt(H), however, tells us the correlation between genotypes and phenotypes in the population, and this permits a probabilistic inference concerning the average amount of difference between individuals obtained IQs and the "genotypic value" of their intelligence.* (The average correlation between phenotypes and genotypes for IQ is about .9 in European and North American Caucasian populations, as determined from summary data presented later in this paper [Table 2]. The square of this value is known as the heritability—the proportion of the phenotypic variance due to genetic variation.) The principle is the same as estimating the "true" scores from obtained scored in test theory. Statements about individuals can be made only on a probabilistic basis and not with absolute certainty. Only if heritability were unity (i.e., H=1) would there be a perfect correlation between obtained scores and genotypic values, in which case we could say with assurance that an individual's measured IQ perfectly represented his genotype for intelligence. This still would not mean that the phenotype could have developed without an environment, for without either heredity or environment there simply is no organism and no phenotype. Thus the statement we so often hear in discussions of individual differences —that the individual's intelligence is the product of the interaction of his heredity and his environment—is rather fatuous. It really states nothing more than the fact that the individual exists.

Constancy. From what has already been said about heritability, it must be clear that it is not a constant like π and the speed of light. H is an empirically determined population statistic, and like any statistic, its value is affected by the characteristics of the population. H will be higher in a population in which environmental variation relevant to the trait in question is small, than in a population in which there is great environmental variation. Similarly, when a population is relatively homogeneous in genetic factors but not in the environmental factors relevant to the development of the characteristic, the heritability of the characteristic will be lower. In short, the value of H is jointly a function of genetic and environmental variability in the population. Also, like any other statistic, it is an estimate based on a sample of the population and is therefore subject to sampling error—the smaller the sample, the greater the margin of probable error. Values of H reported in the literature do not represent what the heritability might be under any environmental conditions or in all populations or even in the same population at different times. Estimates of H are specific to the population sampled, the point in time, how the measurements were made, and the particular test used to obtain the measurements.

Measurement versus Reality. It is frequently argued that since we cannot really measure intelligence we cannot possibly determine its heritability. Whether we can or cannot measure intelligence, which is a separate issue I have already discussed, let it be emphasized that it makes no difference to the question of heritability. We do not estimate the heritability of some trait that lies hidden behind our measurements. We estimate the heritability of the phenotypes and these are the measurements themselves. Regardless of what it is that our tests measure, the heritability tells us how much of the variance in these measurements is due to genetic factors. If the test scores get at nothing genetic, the result will simply be that estimates of their heritability will not differ significantly from zero. The fact that heritability estimates based on IQs differ very significantly from zero is proof that genetic factors play a part in individual differences in IQ. To the extent that a test is not "culture-free" or "culture-fair," it will result in a lower heritability measurement. It makes no more sense to say that intelligence tests do not really measure intelligence but only *developed* intelligence than to say that scales do not really measure a person's weight but only the weight he has acquired by eating. An "environment-free" test of intelligence makes as much sense as a "nutrition-free" scale for weight.

Know All versus Know Nothing. This expression describes another confused notion: the idea that unless we can know absolutely *everything* about the genetics of intelligence we can know nothing! Proponents of this view demand that we be able to spell out in detail every single link in the chain of causality from genes (or DNA molecules) to test scores if we are to say anything about the heritability of intelligence. Determining the heritability of a characteristic does not at all depend upon a knowledge of its physical, biochemical, or physiological basis or of the precise mechanisms through which the characteristic is modified by the environment. Knowledge of these factors is, of course, important in its own right, but we need not have such knowledge to establish the genetic basis of the characteristic. Selective breeding was practiced fruitfully for centuries before anything at all was known of chromosomes and genes, and the science of quantitative genetics upon which the estimation of heritability depends has proven its value independently of advances of biochemical and physiological genetics.

Acquired versus Inherited. How can a socially defined attribute such as intelligence be said to be inherited? Or something that is so obviously acquired from the social environment as vocabulary? Strictly speaking, of course, only genes are inherited. But the brain mechanisms which are involved in learning are genetically conditioned just as are other structures and functions of the organism. What the organism is capable of learning form the environment and its rate of learning thus have a biological basis. Individuals differ markedly in the amount, rate, and kinds of learning they evince even given equal opportunities. Consider the differences that show up when a Mozart and the average run of children are given music lessons! If a test of vocabulary shows high heritability, it only means that persons in the population have had fairly equal opportunity for learning all the words in the test, and the differences in their scores are due mostly to differences in capacity for learning. If members of the population had had very unequal exposures to the words in the vocabulary test, the heritability of the scores would be very low.

Immutability. High heritability by itself does not necessarily imply that the characteristic is immutable. Under greatly changed environmental conditions, the heritability may have some other value, or it may remain the same while the mean of the population changes. At one time tuberculosis had a very high heritability, the reason being that the tuberculosis bacilli were extremely widespread throughout the population, so that the main factor determining whether an individual contracted tuberculosis was not the probability of exposure but the individual's inherited physical constitution. Now that tuberculosis bacilli are relatively rare, difference in exposure rather than in physical predisposition is a more important determinant of who contracts tuberculosis. In the absence of exposure, individual differences in predisposition are of no consequence.

Heritability also tells us something about the locus of control of a characteristic. The control of highly heritable characteristics is usually in the organism's internal biochemical mechanisms. Traits of low heritability are usually controlled by external environmental factors. No amount of psychotherapy, tutoring, or other psychological intervention will elicit normal performance from a child who is mentally retarded because of phenylketonuria (PKU), a recessive genetic defect of metabolism which results in brain damage. Yet a child who has inherited the genes for PKU can grow up normally if his diet is controlled to eliminate certain proteins which contain phenylalanine. Knowledge of the genetic and metabolic basis of this condition in recent years has saved many children from mental retardation.

Parent-Child Resemblance. The old maxim that "like begets like" is held up as an instance of the workings of heredity. The lack of parent-child resemblance, on the other hand, is often mistakenly interpreted as evidence that a characteristic is not highly heritable. But the principles of genetics also explain the fact that often "like begets unlike." A high degree of parent-offspring resemblance, in fact, is to be expected only in highly inbred (or homozygous) strains, as in certain highly selected breeds of dogs and laboratory strains of mice. The random segregation of the parental genes in the formation of the sex cells means that the child receives a random selection of only half of each parent's genes. This fact that parent and child have only 50 percent of their genes in common, along with the effects of dominance and epistasis, insures considerable genetic dissimilarity between parent and child as well as among siblings, who also have only 50 percent of their genes in common. The fact that one parent and a child have only 50 percent of their genes in common is reflected in the average parent-offspring correlation (r_{po}) of between .50 and .60 (depending on the degree of assortative mating for a given characteristic) which obtains for height, head circumference, fingerprint ridges, intelligence, and other highly heritable characteristics. (The resemblance is generally much *higher* than this for traits of *low* heritability.) The genetic correlation between the average of both parents (called the "midparent") and a single offspring ($\underline{r}_{p'o}$) is the square root of the correlation for a single parent (i.e., $\underline{r}_{p'o}) = \text{sqrt}(r_{po})$. The correlation between the average of *both* parents and the average of *all* the offspring ("midchild") that they could theoretically produce ($\underline{r}_{p'o'}$) is the same value as H_N, i.e., heritability in the narrow sense.[4] It is noteworthy that empirical determinations of the midparent-midchild correlation ($\underline{r}_{p'o'}$) in fact closely approximate the values of H as estimated by various methods, such as comparisons of twins, siblings, and unrelated children reared together.

Empirical Findings on the Heritability of Intelligence

It is always preferable, of course, to have estimates of the proportions of variance contributed by each of the components in Equation 2 than to have merely an overall estimate of H. But to obtain reliable estimates of the separate components requires large samples of persons of different kinships, such as identical twins reared together and reared apart, fraternal twins, siblings, half-siblings, parents-children, cousins, and so on. The methods of quantitative genetics by which these variance components, as well as the heritability, can be calculated from such kinship data are technical matters beyond the scope of this article, and the reader must be referred elsewhere for expositions of the methodology of quantitative genetics (Cattell, 1960; Falconer, 1960; Huntley, 1966; Kempthorne, 1957; Loehlin, in press).

The most satisfactory attempt to estimate the separate variance components is the work of Sir Cyril Burt (1955, 1958), based on large samples of many kinships drawn mostly from the school population of London. The IQ test used by Burt was an English adaptation of the Stanford-Binet. Burt's results may be regarded as representative of variance components of intelligence in populations that are similar to the population of London in their degree of genetic heterogeneity and in their range of environment variation. Table 1 shows the percentage of variance due to the various components, groups under "genetic" and "environmental," in Burt's analysis.

TABLE 1
Analysis of Variance of Intelligence Test Scores (Burt, 1958).

Source of Variance	*Percent**	
Genetic:		
Genic (additive)	40.5	(47.9)
Assortative Mating	19.9	(17.9)
Dominance & Epistasis	16.7	(21.7)
Environmental:		
Covariance of Heredity & Environment	10.6	(1.4)
Random Environmental Effects, including HxE interaction (V_1)	5.9	(5.8)
Unreliability (test error)	6.4	(5.3)
Total	100.0	100.0

* Figures in parentheses are percentages for adjusted assessments. See text for explanation.

When Burt submitted the test scores to the children's teachers for criticism on the basis of their impression of the child's "brightness," a number of children were identified for whom the IQ was not a fair estimate of the child's ability in the teacher's judgment. These children were retested, often on a number of tests on several occasions, and the result was an "adjusted" assessment of the child's IQ. The results of the analysis of variance after these adjusted assessments were made are shown in parentheses in Table 1. Note that the component most affected by the adjustments is the covariance of heredity and environment, which is what we should expect if the test is not perfectly "culture-fair." It means that the adjusted scores reduced systematic environmental sources of variance and thereby came closer to representing the children's innate ability, or, stated more technically, the adjusted scores increased the correlation between genotype and phenotype from .88 for the unadjusted scores to .93 for the adjusted scores. [Corrected for test unreliability these correlations become .90 and .96, respectively. And the heritabilities (H_B) for the two sets of scores are therefore $(.90)^2 = .81$ and $(.96)^2 = .93$, respectively.]

Kinship Correlations. the basic data from which variance components and heritability coefficients are estimated are correlations among individuals of different degrees of kinship. Nearly all such kinship correlations reported in the literature are summarized in Table 2. The median values of the correlations obtained in the various studies are given here. These represent the most reliable values we have for the correlation among relatives. Most of the values are taken from the survey by Erlenmeyer-Kimling and Jarvik (1963), and I have supplemented these with certain kinship correlations not included in their survey and reported in the literature since their review (e.g., Burt, 1966, p. 150). The Erlenmeyer-Kimling and Jarvik (1963) review was based on 52 independent studies of the correlation of relatives for tested intellectual abilities, involving over 30,000 correlational pairings from 8 countries in 4 continents, obtained over a period of more than two generations. The correlations were based on a wide variety of mental tests, administered under a variety of conditions by numerous investigators with contrasting views regarding the importance of heredity. The authors conclude: "Against this pronounced heterogeneity, which should have clouded the picture, and is reflected by the wide range of correlations, a clearly definite consistency emerges from the data. The composite data are compatible with the polygenic hypothesis which is generally favored in accounting for inherited differences in mental ability" (Erlenmeyer-Kimling & Jarvik, 1963, p. 1479).

The compatibility with the polygenic hypothesis to which the authors (as outlined earlier on p. 53) refer can be appreciated in Table 2 by comparing the median values of obtained correlations with the sets of theoretical values shown in the last two columns. The first set (Theoretical Value 1) is based on calculations by Burt (1966), using the methods devised by Fisher for estimating kinship correlations for physical characteristics involving assortative mating and some degree of dominance. The second set (Theoretical Value 2) of theoretical values is based on the simplest possible polygenic model, assuming random mating and nothing but additive gene effects. So these are the values one would expect if genetic factors alone were operating and the trait variance reflected no environmental influences whatsoever.

TABLE 2

Correlations for Intellectual Ability: Obtained and Theoretical Values

Correlations Between	*Number of Studies*	*Obtained Median r**	*Theoretical Value 1*	*Theoretical Value 2*
Unrelated Persons				
Children reared apart	4	-.01	.00	.00
Foster parent and child	3	+.20	.00	.00
Children reared together	5	+.24	.00	.00
Collaterals				
Second Cousins	1	+.16	+.14	+.063
First Cousins	3	+.26	+.18	+.125
Uncle (or aunt) and nephew (or niece)	1	+.34	+.31	+.25
Siblings, reared apart	33	+.47	+.52	+.50
Siblings, reared together	36	+.55	+.52	+.50

Dizygotic twins, different sex	9	+.49	+.50	+.50
Dizygotic twins, same sex	11	+.56	+.54	+.50
Monozygotic twins, reared apart	4	+.75	+1.00	+1.00
Monozygotic twins, reared together	14	+.87	+1.00	+1.00
Direct Line				
Grandparent and grandchild	3	+.27	+.31	+.25
Parent (as adult) and child	13	+.50	+.49	+.50
Parent (as child) and child	1	+.56	+.49	+.50

* Correlations not corrected for attenuation (unreliability).
Theoretical Value 1. Assuming assortative mating and partial dominance.
Theoretical Value 2. Assuming random mating and only additive genes, i.e., the simplest possible polygenic model.

First of all, one can note certain systematic departures of the obtained correlations from the theoretical values. These departures are presumably due to non-genetic or environmental influences. The orderly nature of these environmental effects, as reflected in the Erlenmeyer-Kimling and Jarvik median correlations can be highlighted by graphical presentation, as shown in Figure 6. Note that the condition of being reared together or reared apart has the same effect on the difference in magnitudes of the correlation for the various kinships. (The slightly greater difference for unrelated children is probably due to the fact of selective placement by adoption agencies, that is, the attempt to match the child's intelligence with that of the adopting parents.)

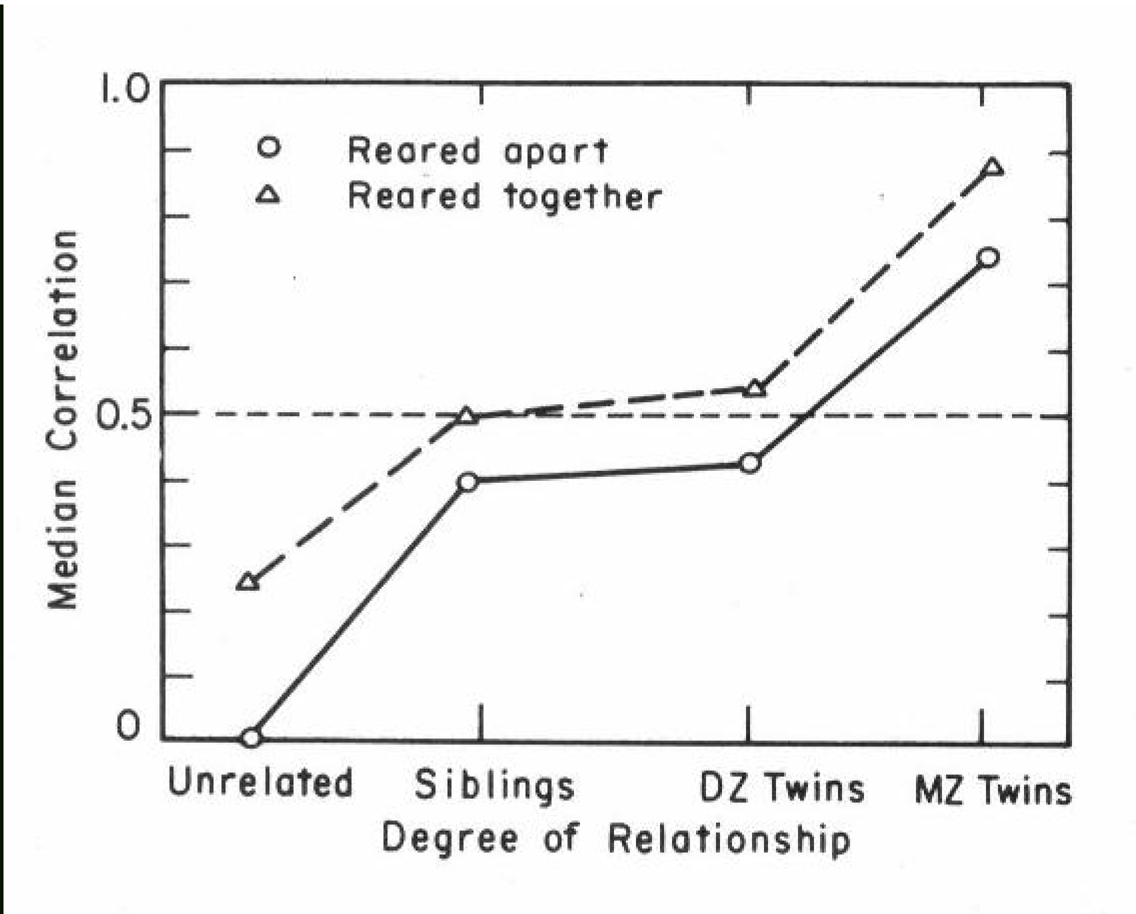

FIGURE 6.
Median values of all correlations reported in the literature up to 1963 for the indicated kinships. (After Erlenmeyer-Kimling & Jarvik, 1963). Note consistency of difference in correlations for relatives reared together and reared apart.

Heritability Estimates. By making certain comparisons among the correlations shown in Table 2 and Figure 6, one can get some insight into how heritability is estimated. For example, we see that the correlation between identical or monozygotic (MZ) twins reared apart is .75. Since MZ twins develop from a single fertilized ovum and thus have exactly the same genes, any difference between the twins must be due to nongenetic factors. And if they are reared apart in uncorrelated environments, the difference between a perfect correlation (1.00) and the obtained correlation (.75) gives an estimate of the proportion of the variance in IQs attributable to environmental differences: 1.00-.75=.25. Thus 75 percent of the variance can be said to be due to genetic variation (this is the heritability) and 25 percent to environmental variation. Now let us go to the other extreme and look at unrelated children reared together. They have no genetic inheritance in common, but they are reared in a common environment. Therefore the correlation between such children will reflect the environment. As seen in Table 2, this correlation is 0.24. Thus, the proportion of IQ variance due to environment is .24; and the remainder, 1.00-.24=.76 is due to heredity. There is quite good agreement between the two estimates of heritability.

Another interesting comparison is between MZ twins reared together (r=.87) and reared apart (r=.75). If 1.00-.75=.25 estimates the total environmental variance, then 1.00-.87=.13 (from MZ twins reared together) is an estimate of the environmental variance *within families* in which children are reared together. Thus the difference between .25-.13=.12 is an estimate of the environmental variance *between families.*

The situation is relatively simple when we deal only with MZ twins, who are genetically identical, or with unrelated children, who have nothing in common genetically. But in order to estimate heritability from any of the other kinship correlations, much more complex formulas are needed which would require much more explanation than is possible in this article. I have presented elsewhere a generalized formula for estimating heritability from any two kinship correlations where one kinship is of a higher degree than the other (Jensen 1967a). I applied this heritability formula to all the correlations for monozygotic and dizygotic (half their genes in common) twins reported in the literature and found an average heritability of .80 for intelligence test scores. (The correlations from which this heritability estimate was derived were corrected for unreliability.) Environmental differences *between* families account for .12 of the total variance, and differences *within* families account for .08. It is possible to derive an overall heritability coefficient from all the kinship correlations given in Table 2. This composite value of H is .77, which becomes .81 after correction for unreliability (assuming an average test reliability of .95). This represents probably the best single overall estimate of the heritability of measured intelligence that we can make. But, as pointed out previously, this in an average value of H about which there is some dispersion of values, depending on such variables as the particular tests used, the population sampled, and the sampling error.

Identical Twins Reared Apart. The conceptually simplest estimate of heritability is, of course, the correlation between identical twins reared apart, since, if their environments are uncorrelated, all they have in common are their genes. The correlation (corrected for unreliability) in this case is the same as the heritability as defined in Equation 3. There have been only three major studies of MZ twins separated early in life and reared apart. All three used individually administered intelligence tests. The correlation between Stanford-Binet IQs of 19 pairs of MZ twins reared apart in a study by Newman, Freeman, and Holzinger (1937) was .77 (.81 corrected for unreliability). The correlation between 44 pairs of MZ twins reared apart on a composite score based on a vocabulary test and Raven's Progressive Matrices was .77 (.81 corrected) in a study by Shields (1962). The correlation between 53 pairs on the Stanford-Binet was .86 (.91 corrected) in a study by Burt (1966). Twin correlations in the same group for height and weight were .94 and .88, respectively.

The Burt study is perhaps the most interesting, for four reasons: (*a*) it is based on the largest sample; (*b*) the IQ distribution of the sample had a mean of 97.8 and a standard deviation of 15.3—values very close to those of the general population; (*c*) all the twin pairs were separated at birth or within their first six months of life; and (*d*) most important, the separated twins were spread over the entire range of socioeconomic levels (based on classification in terms of the six socioeconomic categories of the English census), and there was a slight, though nonsignificant, negative correlation between the environmental ratings of the separated twin pairs. When the twin pairs were rated for differences in the cultural conditions of their rearing, these differences correlated .26 with the differences in their IQs. Differences between material conditions of their homes correlated .16 with IQ differences. (The corresponding correlations for a measure of scholastic attainments were .74 and .37, respectively). The correlation between twins in scholastic attainments was only .62, indicating a much lower heritability than for IQ.)

Foster Parents versus Natural Parents. Children separated from their true parents shortly after birth and reared in adoptive homes show almost the same degree of correlation with the intelligence of their biological parents as do children who are reared by their own parents. The correlations of children with their foster parents' intelligence range between 0 and .20 and are seldom higher than this even when the adoption agency attempts selective placement (e.g., Honzik, 1957). Parent-child correlations gradually increase from zero at 18 months of age to an asymptotic value close to .50 between ages 5 and 6. (Jones, 1954), and this is true whether the child is reared by his parents or not.

Direct Measurement of the Environment. Another method for getting at the relative contribution of environmental factors to IQ variance is simply by correlating children's IQs with ratings of their environment. This can be legitimately done only in the case of adopted children and where there is evidence the selective placement by the adoption agencies is negligible. Without these conditions, of course, some of the correlation between the children and their environmental ratings will be due to genetic factors. There are two large-scale studies in the literature which meet these criteria. Also, both studies involved adopting parents who are representative of a broad cross section of the U.S. Caucasian population with respect to education, occupation, and socioeconomic level. It is probably safe to say that not more than five percent of the U.S. Caucasian population falls outside the range of environmental variation represented in the samples of these two studies. The study by Leahy (1935) found an average correlation of .20 between the IQs of adopted children and a number of indices of the "goodness" of their environment, including the IQs and education of both adoptive parents, their socioeconomic status, and the cultural amenities in the home. Leahy concluded from this that the environment ratings accounted for a 4 percent (i.e., the square of r=.20) of the variance in the adopted children's Stanford-Binet IQs, and that 96 percent of the variance remained to be accounted for by other factors. The main criticisms we can make of this study are, first, that the environmental indices were not sufficiently "fine-grained" to register the subtleties of environmental variation and of the qualities of parent-child relationship that influence intellectual development, and, second, that the study did not make use of the technique of multiple correlation, which would show the total contribution to the variance of all the separate environmental indices simultaneously. A multiple correlation is usually considerably greater than merely the average of all the correlations for the single variables.

A study by Burks' (1928) meets both these objections. To the best of my knowledge, no study before or since has rated environments in any more detailed and fine-grained manner than did Burks'. Each adoptive home was give 4 to 8 hours of individual investigation. As in Leahy's study, Burks' included intelligence measures on the adopting parents as part of the children's environment, an environment which also included such factors as the amount of time the parents spent helping the children with their school work, the amount of time spent reading to the children, and so on. The multiple correlation (corrected for unreliability) between Burks' various environmental ratings and the adopted children's Stanford-Binet IQs was .42. The square of this correlation is .18, which represents the proportion of IQ variance accounted for by Burks' environmental measurements. This value comes very close to the environmental variance estimated in direct heritability analyses based on kinship correlations.

Burks translated her findings into the conclusion that the total effect of environmental factors one standard deviation up or down the environmental scale is only about 6 IQ points. This is an interesting figure, since it is exactly half the 12 point IQ difference found on the average between normal siblings reared together by their own parents. Siblings differ genetically, of course, having only about half their genes in common. If all the siblings in every family were divided into two groups—those above and those below the family average—the IQ distributions of the two groups would appear as shown in Figure 7. Though the average difference is only 12 IQ points, note the implications in the proportions of each group falling into the upper and lower ranges of the IQ scale. It would be most instructive to study the educational and occupational attainments of these two groups, since presumably they should have about the same environmental advantages.

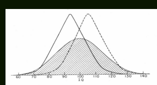

FIGURE 7.
IQ distributions of siblings who are below (solid curve) or above (dashed curve) their family average. The shaded curve is the IQ distribution of randomly selected children.

Another part of Burks' study consisted of a perfectly matched control group of parents rearing their own children, for whom parent-child correlations were obtained. Sewall Wright (1931) performed a heritability analysis on these parent-child and IQ-environment correlations and obtained a heritability coefficient of .81.

Effects of Inbreeding on Intelligence

One of the most impressive lines of evidence for the involvement of genetic factors in intelligence comes from study of the effects of inbreeding, that is, the mating of relatives. In the case of polygenic characteristics the direction of the effect of inbreeding is predictable from purely genetic considerations. All individuals carry in their chromosomes a number of mutant or defective genes. These genes are almost always recessive, so they have no effect on the phenotype unless by rare chance they match up with another mutant gene at the same locus on a homologous chromosome; in other words, the recessive mutant gene at a given locus must be inherited from both the father and mother in order to affect the phenotype. Since such mutants are usually defective, they do not enhance the phenotypic expression of the characteristic but usually degrade it. And for polygenic characteristics we would expect such mutants to lower the metric value of the characteristics by graded amounts, depending upon the number of paired mutant recessives. If the parents are genetically related, there is a greatly increased probability that the mutant recessives at a given loci will be paired in the offspring. The situation is illustrated in Figure 8, which depicts in a simplified way a pair of homologous chromosomes inherited by an individual from a mother (M) and father (F) who are related (Pair A) and a pair of chromosomes inherited from unrelated parents (Pair B). The blackened spaces represent recessives. Although both pairs contain equal numbers of recessives, more of them are at the same loci in Pair A than in Pair B. Only the paired genes degrade the characteristics' phenotypic value.

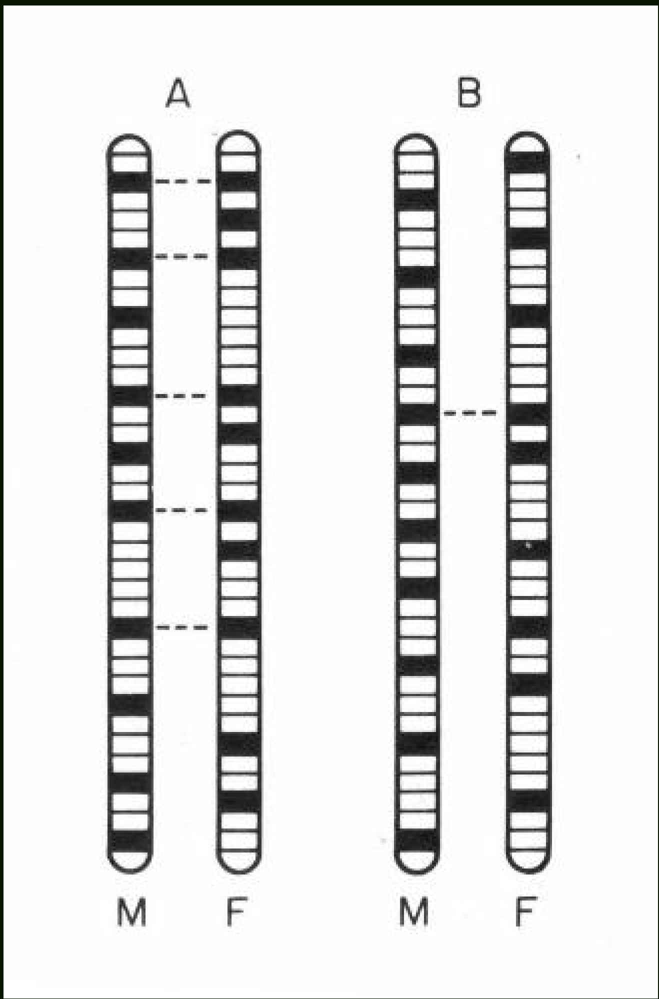

FIGURE 8.
Simplified schema of chromosomes, illustrating the pairing of recessive (mutant) genes (black spaces) in homologous chromosomes from mother (M) and father (F). Pair A has five pairs of recessives in the same loci on the chromosome, Pair B has only one such pair.

A most valuable study of this genetic phenomenon with respect to intelligence was carried out in Japan after World War II by Schull and Neel (1965). The study illustrates how strictly sociological factors, such as mate selection, can have extremely important genetic consequences. In Japan approximately five percent of all marriages are between cousins. Schull and Neel studied the offspring of marriages of first cousins, first cousins once removed, and second cousins. The parents were statistically matched with a control group of unrelated parents for age and socioeconomic factors. Children from the cousin marriages and the control children from unrelated parents (total N=2111) were given the Japanese version of the Wechsler Intelligence Scale for Children (WISC). The degree of consanguinity represented by the cousin marriages in this study had the effect of depressing WISC IQs by an average of 7.4 percent, making the mean of the inbred group nearly 8 IQ points lower than the mean of the control group. Assuming normal distributions of IQ, the effect is shown in Figure 9, and illustrates the point that the most drastic consequences of group mean differences are to be seen in the tails of the distributions. In the same study, a similar depressing effect was found for other polygenic characteristics such as several anthropometric and dental variables.

The mating of relatives closer than cousins can produce a markedly greater reduction in offspring's IQs. Lindzey (1967) has reported that almost half of a group of children born to so-called nuclear incest matings (brother-sister or father-daughter) could not be placed for adoption because of mental retardation and other severe defects which had a relatively low incidence among the offspring of unrelated parents who were matched with the incestuous parents in intelligence, socioeconomic status, age, weight, and stature. In any geographically confined population where social or legal regulations on mating are lax, where individuals' paternity is often dubious, and where the proportion of half-siblings within the same age groups is high, we would expect more inadvertent inbreeding, with its unfavorable genetic consequences, than in a population in which these conditions exist to a lesser degree.

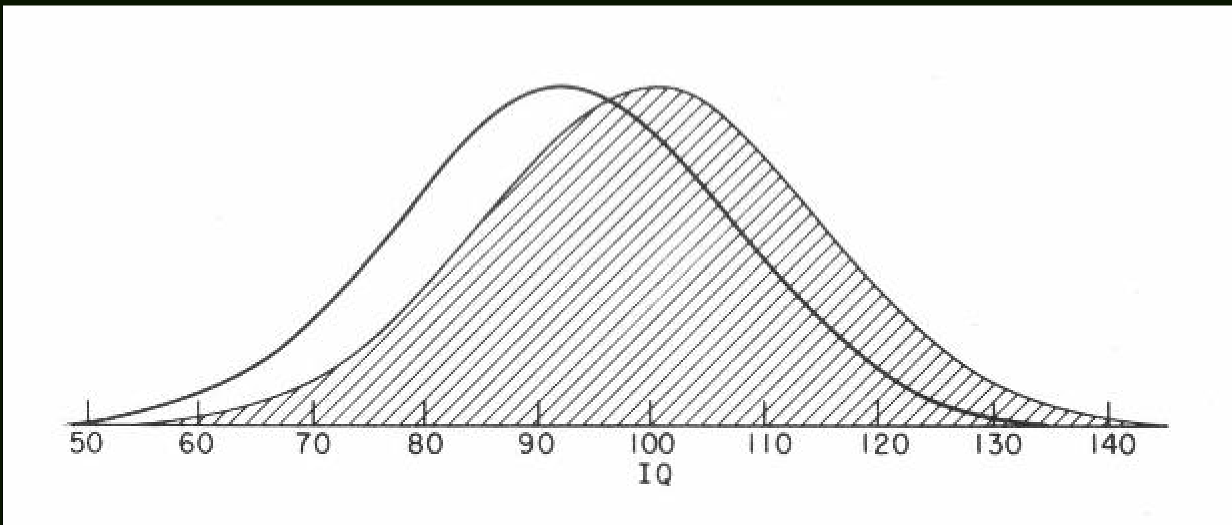

FIGURE 9.
The average effect of inbreeding to the degree of 1st, 1 1/2, and 2nd cousin matings on the IQ distribution of offspring (heavy line). Shaded curve is the IQ distribution of the offspring of nonconsanguinous matings. (After Schull & Neel, 1965.)

Heritability of Special Mental Abilities. When the general factor, or *g*, is removed from a variety of mental tests, the remaining variance is attributable to a number of so-called "group factors" or "special abilities." The tests of special abilities that have been studied most thoroughly with respect to their heritability are Thurstone's Primary Mental Abilities: Verbal, Space, Number, Word Fluency, Memory, and Perceptual Speed. Vandenberg (1967) has reviewed the heritability studies of these tests and reports that the H values range from near zero to about .75, with most of values of H between .50 and .70. Vandenberg devised a method for estimating the genetic components of these special abilities which are completely independent of *g*. He concluded that at least four of the Primary Mental Abilities (Number, Verbal, Space, and Word Fluency) independently have significant hereditary components.

There have been few studies of the heritability of noncognitive skills, but a study by McNemar (see Bilodeau, 1966, Ch. 3) of motor skill learning indicates that heritabilities in this sphere may be even higher than for intelligence. The motor skill learning was measured with a pursuit-rotor, a tracking task in which the subject must learn to keep a stylus on a metal disc about the size of a nickel rotating through a circumference of about 36 inches at 60 rpm. The percentage of time "on target" during the course of practice yields a learning measure of high reliability, showing marked individual differences both in rate of acquisition and final asymptote of this perceptual motor skill. Identical twins correlated .95 and fraternal twins .51 on pursuit-rotor learning, yielding a heritability coefficient of .88, which is very close to the heritability of physical stature.

Heritability of Scholastic Achievement. The heritability of measures of scholastic achievement is much less, on the average, than the heritability of intelligence. In reviewing all the twin studies in the literature containing relevant data, I concluded that individual differences in scholastic performance are determined less than half as much by heredity as are differences in intelligence (Jensen, 1967a).[5] The analysis of all the twin studies on a variety of scholastic measures gives an average H of .40. The environmental variance of 60 percent can be partitioned into variance due to environmental differences *between* families, which is 54 percent, and the differences *within* families is 6 percent. But it should also be noted that the heritability estimates for scholastic achievement vary over a much wider range than do H values for intelligence. In general, H for scholastic achievement increases as we go from the primary grades up to high school and it is somewhat lower for relatively simple forms of learning (e.g., spelling and arithmetic computation) than for more complex learning (e.g., reading comprehension and arithmetic problem solving). Yet large-sample twin data from the National Merit Scholarship Corporation show that the *between families* environmental component accounts for about 60 percent of the variance in students' rank in their high school graduating class. This must mean that there are strong family influences which cause children to conform to some academic standard set by the family and which reduce variance in scholastic performance among siblings reared in the same family. Unrelated children reared together are also much more alike in school performance than in intelligence. The common finding of a negative correlation between children's IQ and the amount of time parents report spending in helping their children with school work is further evidence that considerable family pressures are exerted to equalize the scholastic performance of siblings. This pressure to conform to a family standard shows up most conspicuously in the small *within families* environmental variance component on those school subject which are most susceptible to improvement by extra coaching, such as spelling and arithmetic computation.

The fact that scholastic achievement is considerably less heritable than intelligence also means that many other traits, habits, attitudes, and values enter into a child's performance in school besides just his intelligence, and these non-cognitive factors are largely environmentally determined, mainly through influences within the child's family. This means there is potentially much more we can do to improve school performance through environmental means than we can do to change intelligence per se. Thus it seems likely that if compensatory education programs are to have a beneficial effect on achievement, it will be through their influence on motivation, values, and other environmentally conditioned habits that play and important part in scholastic performance, rather than through any marked direct influence on intelligence per se. The proper evaluation of such programs should therefore be sought in their effects on actual scholastic performance rather than in how much their raise the child's IQ.

How the Environment Works

Environment as a Threshold

All the reports I have found of especially large upward shifts in IQ which are explicitly associated with environmental factors have involved young children, usually under six years of age, whose initial social environment was deplorable to a greater extreme than can be found among any children who are free to interact with other persons or to run about out-of-doors. There can be no doubt that moving children from an extremely deprived environment to good average environmental circumstances can boost the IQ some 20 to 30 points and in certain extreme rare cases as much as 60 or 70 points. On the other hand, children reared in rather average circumstances do not show an appreciable IQ gain as a result of being placed in a more culturally enriched environment. While there are reports of groups of children going from below average up to average IQs as a result of environmental enrichment, I have found no report of a group of children being given permanently superior IQs by means of environmental manipulations. In brief, it is doubtful that psychologists have found consistent evidence for any social environmental influences short of extreme environmental isolation which have a marked systematic effect on intelligence. This suggests that the influence of the quality of the environment on intellectual development is not a linear function. Below a certain threshold of environmental adequacy, deprivation can have a markedly depressing effect on intelligence. But above this threshold, environmental variations cause relatively small differences in intelligence. The fact that the vast majority of the populations sampled in studies of the heritability of intelligence are above this threshold level of environmental adequacy accounts for the high values of the heritability estimates and the relatively small proportion of IQ variance attributable to environmental influences.

The environment with respect to intelligence is thus analogous to nutrition with respect to stature. If there are great nutritional lacks, growth is stunted, but above a certain level of nutritional adequacy, including minimal daily requirements of minerals, vitamins, and proteins, even great variations in eating habits will have negligible effects on persons' stature, and under such conditions most of the differences in stature among individuals will be due to heredity.

When I speak of subthreshold environmental deprivation, I do not refer to a mere lack of middle-class amenities. I refer to the extreme sensory and motor restrictions in environments such as those described by Skeels and Dye (1939) and Davis (1947), in which the subjects had little sensory stimulation of any kind and little contact with adults. These cases of extreme social isolation early in life showed great deficiencies in IQ. But removal from social deprivation to a good, average social environment resulted in large gains in IQ. The Skeels and Dye orphanage children gained in IQ from an average of 64 at 19 months of age to 96 at age 6 as a result of being given social stimulation and placement in good homes between 2 and 3 years of age. When these children were followed up as adults, they were found to be average citizens in their communities, and their own children had an average IQ of 105 and were doing satisfactorily in school. A far more extreme case was that of Isabel, a child who was confined and reared in an attic up to the age of six by a deaf-mute mother, and who had an IQ of about 30 at age 6. When Isabel was put into a good environment at that age, her IQ became normal by age 8 and she was able to perform as an average student throughout school (Davis, 1947). Extreme environmental deprivation thus need not permanently result in below average intelligence.

These observations are consistent with studies of the effects of extreme sensory deprivation on primates. Monkeys raised from birth under conditions of total social isolation, for example, show no indication when compared with normally raised controls, of any permanent impairment of ability for complex discrimination learning, delayed response learning, or learning set formation, although the isolated monkeys show severe social impairment in their relationships to normally reared monkeys (Harlow & Griffin, 1965).

Thoughtful scrutiny of all these studies of extreme environmental deprivation leads to two observations which are rarely made by psychologists who cite the studies as illustrative explanations of the low IQs and poor scholastic performance of the many children called culturally disadvantaged. In the first place, typical culturally disadvantaged children are not reared in anything like the degree of sensory and motor deprivation that characterizes, say, the children of the Skeels study. Secondly, the IQs of severely deprived children are markedly depressed even at a very early age, and when they are later exposed to normal environmental stimulation, their IQs rise rapidly, markedly, and permanently. Children called culturally disadvantaged, on the other hand, generally show no early deficit and are usually average and sometimes precocious on perceptual-motor tests administered before two years of age. The orphanage children described in Skeels' study are in striking contrast to typical culturally disadvantaged children of the same age. Also, culturally disadvantaged children usually show a slight initial gain in EQ after their first few months of exposure to the environmental enrichment afforded by school attendance, but, unlike Skeels' orphans, they soon lose this gain, and in a sizeable proportion of children the initial IQ gain is followed by a gradual decline in IQ throughout the subsequent years of schooling. We do not now how much of this decline is related to environmental or hereditary factors. We do know that with increasing age children's IQs increasingly resemble their parents' rank order in intelligence whether they are reared by them or not, and therefore with increasing age we should expect greater and more reliable differentiation among children's IQs as they gravitate toward their genotypic values (Honzik, 1957). Of course, the gravitating effect is compounded by the fact that less intelligent parents are also less apt to provide the environmental conditions conducive to intellectual development in the important period between ages 3 and 7, during which children normally gain increasing verbal control over their environment and their own behavior. (I have described some of these environmental factors in detail elsewhere [Jensen, 1968e].)

Heber, Dever, and Conry (1968) have obtained data which illustrate this phenomenon of children's gravitation toward the parental IQ with increasing age. They studied the families of 88 low economic class Negro mothers residing in Milwaukee in a set of contiguous slum census tracts, an area which yields the highest known prevalence of identified retardation in the city's schools. Although these tracts contribute about 5 percent of the schools' population, they account for about one-third of the school children classed as mentally retarded (IQ below 75). The sample of 88 mothers was selected by taking 88 consecutive births in these tracts where the mother already had at least one child of age six. The 88 mothers had a total of 586 children, excluding their newborns. The percentage of mothers with IQs of 80 or above 54.6; 45.4 percent were below IQ 80. The IQs of the children of these two groups of mothers were plotted as a function of the children's age. The results are shown in Figure 10. Note that only the children whose mothers' IQs are below 80 show a systematic decline in IQ as well as a short-lived spurt of several points at the age of entrance into school. At six years of age and older, 80.8 percent of the children with IQ's below 80 were those whose mothers had IQs below 80.

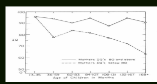

FIGURE 10.
Mean IQs of 586 children of 88 mothers as a function of age of children. (Heber, Dever, & Conry, 1968.)

It is far from certain or even likely that all such decline in IQ is due to environmental influences rather than to genetic factors involved in the growth rate of intelligence. Consistent with this interpretation is the fact that the heritability of intelligence measures increases with age. We should expect just the opposite if environmental factors alone were responsible for the increasing IQ deficit of markedly below average groups. A study by Wheeler (1942) suggests that although IQ may be raised at all age levels by improving the environment, such improvements do not counteract the decline in the IQ of certain below-average groups. In 1940 Wheeler tested over 3000 Tennessee mountain children between the ages of 6 and 16 and compared their IQs with children in the same age range who had been given the same tests in 1930, when the average IQ and standard of living in this area would characterize the majority of the inhabitants as "culturally deprived." During the intervening 10 years state and federal intervention in this area brought about great improvements in economic conditions, standards of health care, and educational and cultural opportunities, and during the same period the average IQ for the region increased 10 points, from 82 to 92. But the decline in IQ from age 6 to age 16 was about the same in 1940 (from 103 to 80) as in 1930 (from 95 to 74).

Reaction Range. Geneticists refer to the concept of reaction range (RR) in discussing the fact that similar genotypes may result in quite different phenotypes depending on the favorableness of the environment for the development of the characteristic in question. Of further interest to geneticists is the fact that different genotypes may have quite different reaction ranges; some genotypes may be more buffered against environmental influences than others. Different genetic strains can be unequal in their susceptibility to the same range of environmental variation, and when this is the case, the strains will show dissimilar heritabilities on the trait in question, the dissimilarity being accentuated by increasing environmental variation. Both of these aspects of the reaction range concept are illustrated hypothetically with respect to IQ in Figure 11.

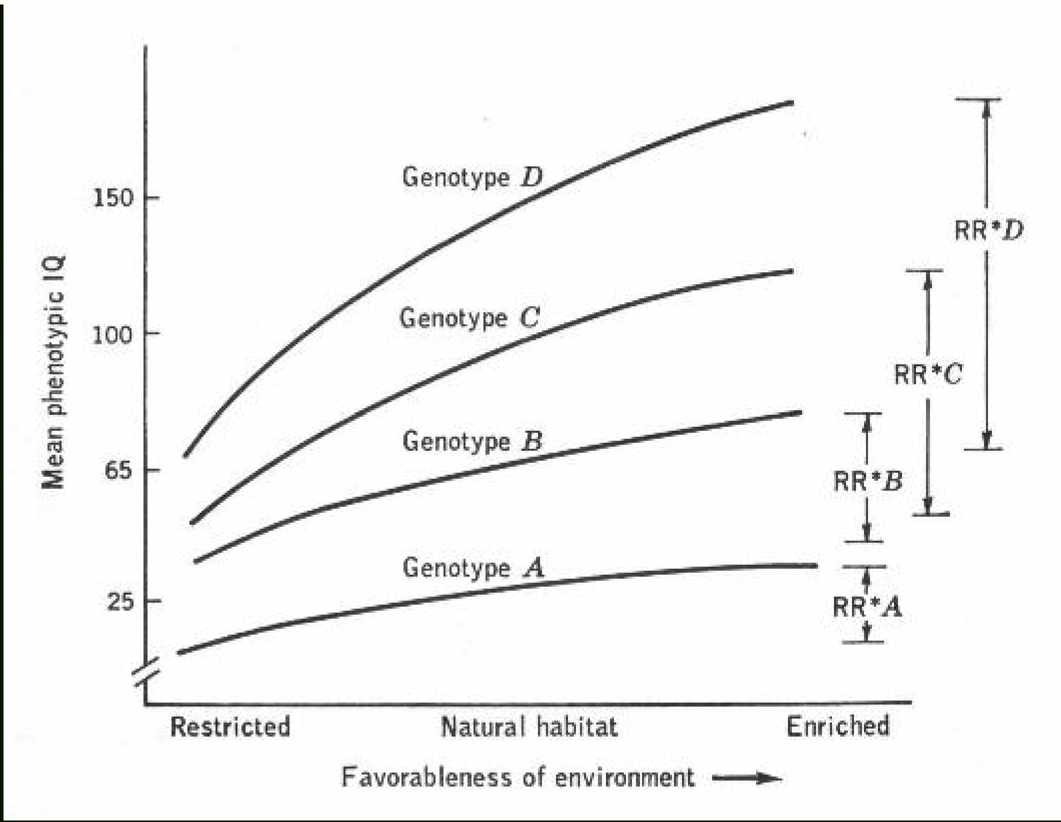

FIGURE 11.
Scheme of the reaction range concept for four hypothetical genotypes. RR denotes the presumed reaction range for phenotypic IQ. Note: Large deviations from the "natural habitat" have a low probability of occurrence. (From Gottesman, 1963.)

The above discussion should serve to counter a common misunderstanding about quantitative estimates of heritability. It is sometimes forgotten that such estimates actually represent *average* values in the population that has been sampled and they do not necessarily apply either to differences *within* various subpopulations or to differences *between* subpopulations. In a population in which an overall H estimate is, say, .80, we may find a certain group for which H is only .70 and another group for which H is .90. All the major heritability studies reported in the literature are based on samples of White European and North American populations, and our knowledge of the heritability of intelligence in different racial and cultural groups within these populations is nil. For example, no adequate heritability studies have been based on samples of the Negro population of the United States. Since some genetic strains may be more buffered from environmental influences than others, it is not sufficient merely to equate the environments of various subgroups in the population to infer equal heritability of some characteristic in all of them. The question of whether heritability estimates can contribute anything to our understanding of the relative importance of genetic and environmental factors in accounting for average phenotypic differences between racial groups (or any other socially identifiable groups) is too complex to be considered here. I have discussed this problem in detail elsewhere and concluded that heritability estimates could be of value in testing certain specific hypotheses in this area of inquiry, provided certain conditions were met and certain other crucial items of information were also available (Jensen, 1968c).

Before continuing discussion of environmental factors we must guard against one other misunderstanding about heritability that sometimes creeps in at this point. This is the notion that because so many different environmental factors and all their interactions influence the development of intelligence, by the time the child is old enough to be tested, these influences must totally bury or obscure all traces of genetic factors —the genotype must lie hidden and inaccessible under the heavy overlay of environmental influences. If this were so, of course, the obtained values of H would be very close to zero. But the fact that values of H for intelligence are usually quite high (in the region of .70 to .90) means that current intelligences tests can, so to speak, "read through" the environmental "overlay."

Physical versus Social Environment

The value 1-H, which for IQ generally amounts to about .20, can be called E, the proportion of variance due to nongenetic factors. There has been a pronounced tendency to think of E as being wholly associated with individuals' social and interpersonal environment, child rearing practices, and differences in educational and cultural opportunities afforded by socioeconomic status. It is certain, however, that these sociological factors are not responsible for the whole E and it is not improbable that they contribute only a minor portion of the E variance in the bulk of our population. Certain physical and biological environmental factors may be at least as important as the social factors in determining individual differences in intelligence. If this is true, advances in medicine, nutrition, prenatal care, and obstetrics may contribute as much or more to improving intelligence as will manipulation of the social environments.

Prenatal Environment of Twins. A little known fact about twins is that they average some 4 to 7 points lower in IQ than singletons (Vandenberg, 1968). The difference shows up in scholastic achievement, as shown in the distribution of reading scores of twin and singleton girls in Sweden (Figure 12).

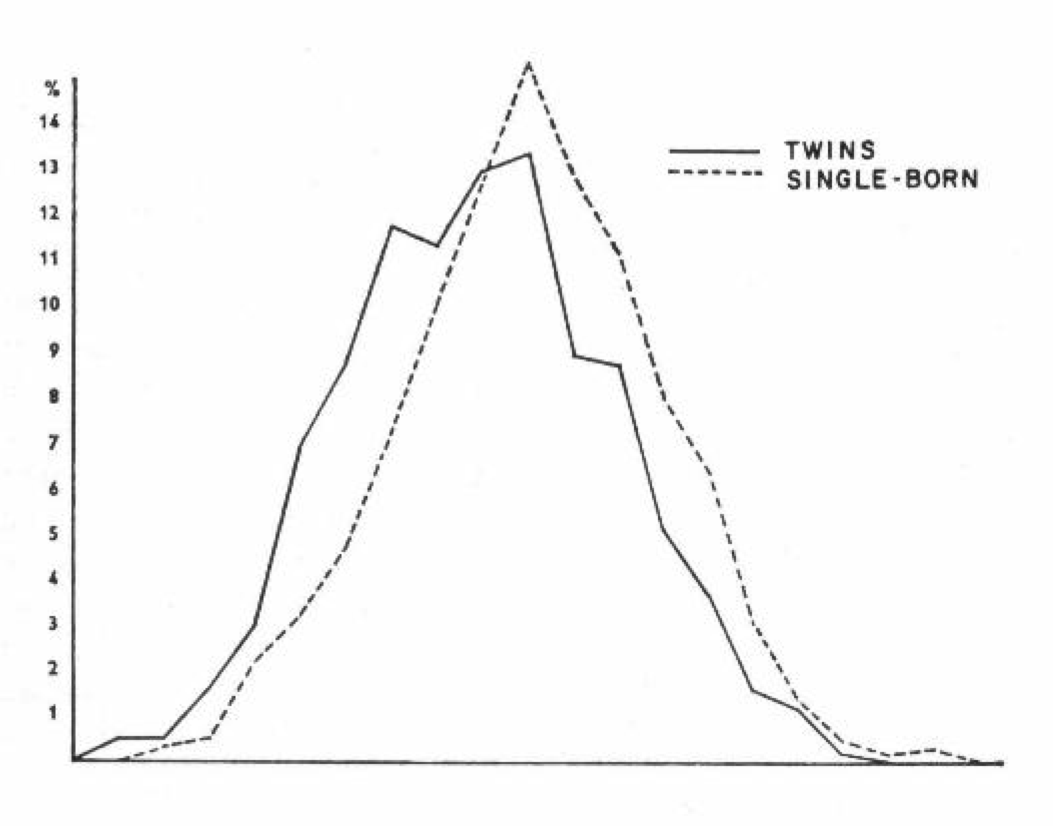

FIGURE 12.
Distribution of reading scores of twins and single children (all girls). (Husén, 1960.)

If this phenomenon were due entirely to differences between twins and singletons in the amount of individual attention they receive from heir parents, one might expect the twin-singleton difference to be related to the family's socioeconomic status. But there seems to be no systematic relationship of this kind. The largest study of the question, summarized in Figure 13, shows about the same average amount of twin-singleton IQ disparity over a wide range of socioeconomic groups.

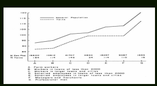

FIGURE 13.
Distribution of IQs by occupation of father, for twins and singletons. (Zasso, 1960.)

Three other lines of evidence place the locus of this effect in the prenatal environment. Monozygotic twins are slightly lower in IQ than dizygotic twins (Stott, 1960, p. 98), a fact which is consistent with the finding that MZ twins have a higher mortality rater and greater disparity in birth weights than DZ twins, suggesting that MZ twins enjoy less equal and less optimal intrauterine conditions than DZ twins or singletons. Inequalities in both intrauterine space and fetal nutrition probably account for this. Also, boy twins are significantly lower in IQ than girl twins, which conforms to the well known greater vulnerability of male infants to prenatal impairment (Stott, 1960). Finally, the birth weight of infants, when matched for gestational age, is slightly but significantly correlated with later IQ, and the effect is independent of sociocultural factors (Churchill, Neff, & Caldwell, 1966). In pairs of identical twins, the twin with the lower birth weight usually has the lower IQ (by 5 to 7 points on the average) at school age. This is true both in white and Negro twins. The birth-weight differences are reflected in all 11 subtests of the Wechsler Intelligence Scale for Children and are slightly greater on the Performance than on the Verbal tests (Willerman & Churchill, 1967). The investigators interpret these findings as suggesting that nutrient supplies may be inadequate for proper body and brain development in twin pregnancies, and that unequal sharing of nutrients and space stunts one twin more than its mate.

Thus, much of the average difference between MZ twins, whether reared together or reared apart, seems to be due to prenatal environmental factors. The real importance of these findings, of course, lies in their implications for the possible role of prenatal environment in the development of all children. It is not unlikely that there are individual maternal differences in the adequacy of the prenatal environment. If intrauterine conditions can cause several points of IQ difference between twins, it is not hard to imagine that individual differences in prenatal environments could also cause IQ differences in single born children and might therefore account for a substantial proportion of the total environmental variance in IQ.

Abdominal Decompression. There is now evidence that certain manipulations of the intrauterine environment can affect the infant's behavioral development for many months after birth. A technique known as abdominal decompression was invented by a professor of obstetrics (Heyns, 1963), originally for the purpose of making women experience less discomfort in the latter months of their pregnancy and also to facilitate labor and delivery. For about an hour a day during the last three or four months of pregnancy, the woman is placed in a device that creates a partial vacuum around her abdomen, which greatly relaxes the intrauterine pressure. The device is used during labor up to the moment of delivery. Heyns has applied this device to more than 400 women. Their infants as compared with control groups who have not received this treatment, show more rapid development in their first two years and manifest an overall superiority in tests of perceptual-motor development. They sit up earlier, walk earlier, talk earlier, and appear generally more precocious than their own siblings or other children whose mothers were not so treated. At two years of age, the children in Heyns' experiment had DQs (developmental quotients) some 30 points higher than the control children (in the general population the mean DQ is 100, with a standard deviation of 15). Heyns explains the effect of maternal abdominal decompression on the child's early development in terms of the reduction of intrauterine pressure, which results in a more optimal blood supply to the fetus and also lessens the chances of brain damage during labor. (The intrauterine pressure on the infant's head is reduced from about 22 pounds to 8 pounds.) Results of children's later IQs have not been published, but correspondence with Professor Heyns and verbal reports from visitors to his laboratory inform me that there is no evidence that the IQ of these children is appreciably higher beyond age 6 than that of control groups. If this observation is confirmed by the proper methods, it should not be too surprising in view of the negligible correlations normally found between DQs and later IQs. But since abdominal decompression results in infant precocity, one may wonder to what extent differences in intrauterine pressure are responsible for normal individual and group differences in infant precocity. Negro infants, for example, are more precocious in development (as measured on the Bayley Scales) in their first year or two than Caucasian infants (Bayley, 1965a). Infant precocity would seem to be associated with more optimal intrauterine and perinatal conditions. This conjecture is consistent with the finding that infants whose prenatal and perinatal histories would make them suspect of some degree of brain damage show lower DQs on the Bayley Scales than normal infants (Honzik, 1962). Writers who place great emphasis on the hypothesis of inadequate prenatal care and complications of pregnancy to account for the lower average IQ of Negroes (e.g., Bronfenbrenner, 1967) are also obliged to explain why these unfavorable factors do not also depress the DQ below average in Negro infants, as do such factors as brain damage and prenatal and infant malnutrition (Craviotos, 1966). Since all such environmental factors should lower the heritability of intelligence in any segment of the population in which they are hypothesized to play an especially significant role, one way to test the hypothesis would be to compare the heritability of intelligence in that segment of the population for which extra environmental factors are hypothesized with the heritability in other groups for whom environmental factors are supposedly less accountable for IQ variance.

A Continuum of Reproductive Casualty. A host of conditions associated with reproduction which are known to differ greatly across socioeconomic levels have been hypothesized as causal factors in average intellectual differences. There is no doubt about the fact of the greater prevalence in poverty areas of conditions unfavorable to optimal pregnancy and safe delivery. The question that remains unanswered is the amount of IQ variance associated with these conditions predisposing to reproductive casualty. The disadvantageous factors most highly associated with social conditions are: pregnancies at early ages, teenage deliveries, pregnancies in close succession, a large number of pregnancies, and pregnancies that occur late in the woman's reproductive life (Graves, Freeman, & Thompson, 1968). These conditions are related to low birth weight, prematurity, increased infant mortality, prolonged labor, toxemia, anemia, malformations, and mental deficiency in the offspring. Since all of these factors have a higher incidence in low socioeconomic groups and in certain ethnic groups (Negroes, American Indians, and Mexican-Americans) in the United States, the probably account for some proportion of the group differences in IQ and scholastic performance, but just how much of the true differences they may account for no one really knows at present. It is interesting that Jewish immigrants, whose offspring are usually found to have a higher mean IQ than the general population, show fewer disadvantageous reproductive conditions and have the lowest infant mortality rates of all ethnic groups, even when matched with other immigrant and native born groups on general environmental conditions (Graves et al., 1968).

Although disadvantageous reproductive factors occur differentially in different segments of the population, it is not at all certain how much they are responsible for the IQ differences between social classes and races. It is reported by the National Institute of Neurological Diseases and Blindness, for example, that when all cases of mental retardation that can be reasonably explained in terms of known complications of pregnancy and delivery, brain damage, or major gene and chromosomal defects are accounted for, there still remain 75 to 80 percent of the cases who show no such specific causes and presumably represent just the lower end of the normal polygenic distribution of intelligence (Research Profile No. 11, 1965). Buck (1968) has argued that it still remains to be proven that a degree of neurological damage is bound to occur among the survivors of all situations which carry a high risk of perinatal mortality and that a high or even a known proportion of mental retardation can be ascribed to the non-lethal grades of reproductive difficulty. A large study reported by Buck (1968) indicates that the most common reproductive difficulties when occurring singly have no significant effect on children's intellectual status after age 5, with one exception of pre-eclamptic toxemia of pregnancy, which caused some cognitive impairment. Most of the complications of pregnancy, it seems, must occur multiply to impair intellectual ability. It is as if the nervous system is sufficiently homeostatic to withstand certain unfavorable conditions if they occur singly.

Prematurity. The literature on the relationship of premature birth to the child's IQ is confusing and conflicting. Guilford (1967), in his recent book on *The Nature of Intelligence*, for example, concluded, as did Stoddard (1943), that prematurity has no effect on intelligence. Stott (1966), on the other hand, presents impressive evidence of very significant IQ decrements associated with prematurity. Probably the most thorough review of the subject I have found, by Kushlick (1966), helps to resolve these conflicting opinions. There is little question that prematurity has the strongest known relation to brain dysfunction of any reproductive factor, and many of the complications of pregnancy are strongly associated with the production of premature children. The crucial factor in prematurity, however, is not prematurity per se, but low birth-weight. Birth-weight apparently acts as a threshold variable with respect to intellectual impairment. All studies of birth-weight agree in showing that the incidence of babies weighing less than 5

1/2 lbs. increases from higher to lower social classes. But only about 1 percent of the total variance of birth-weight is accounted for by socioeconomic variables. Race (Negro versus white) has an effect on birth-weight independently of socioeconomic variables. Negro babies mature at a lower birth-weight than white babies (Naylor & Myrianthopoulos, 1967). If prematurity is defined as a condition in which birth-weight is under 5 1/2 lbs., the observed relationship between prematurity and depression of the IQ is due to the common factor of low social class. Kushlick (1966, p. 143) concludes that it is only among children having birth-weights under 3 lbs. that the mean IQ is lowered, independently of social class, and more in boys than in girls. The incidence of extreme subnormality is higher for children with birth-weights under 3 or 4 lbs. But when one does not count these extreme cases (IQs below 50), the effects of prematurity or low birth-weight—even as low as 3 lbs.—have a very weak relationship to children's IQs by the time they are of school age. The association between very low birth-weight and extreme mental subnormality raises the question of whether the low birth-weight causes the abnormality or whether the abnormality arises independently and causes the low birth-weight.

Prematurity and low birth-weight have a markedly higher incidence among Negroes than among whites. That birth-weight differences per se are not a predominant factor in Negro-white IQ differences, however, is suggested by the findings of a study which compared Negro and white premature children matched for birth-weight. The Negro children in all weight groups performed significantly less well on mental tests at 3 and 5 years of age than the white children of comparable birth-weight (Hardy, 1965, p. 51).

Genetic Predisposition to Prenatal Impairment. Dennis Stott (1960, 1966), a British psychologist, has adduced considerable evidence for the theory that impairments of the central nervous system occurring prenatally as a result of various stresses in pregnancy may not be the *direct* result of adverse intrauterine factors but may result *indirectly* from genetically determined mechanisms which are triggered by prenatal stress of one form or another.

Why should there exist a genetic mechanism predisposing to congenital impairments? Would not such genes, if they had ever existed, have been eliminated long ago through natural selection? It can be argued from considerable evidence in lower species of mammal observable by zoologists today that such a genetic mechanism may have had survival value for primitive man, but that the correlations of our present industrial society and advances in medical care have diminished the biological advantage of this mechanism for survival of the human species. The argument is that, because of the need to control population, there is a genetic provision within all species for multiple impairments, which are normally only potentialities, that can be triggered off by prenatal stress associated with high population density, such as malnutrition, fatigue from overexertion, emotion distress, infections, and the like. The resulting congenital impairment would tend to cut down the infant population, thereby relieving the pressure of population without appreciably reducing the functioning and efficiency of the young adults in the population. Stott (1966) has presented direct evidence of an association between stresses in the mother during pregnancy and later behavioral abnormalities and learning problems of the child in school. The imperfect correlation between such prenatal stress factors and signs of congenital impairment suggests that there are individual differences in genetic predisposition to prenatal impairment. The hypothesis warrants further investigation. The prenatal environment could be a much more important source of later IQ variance for some children than for others.

Mother-Child Rh Incompatibility. The *Rh* blood factor can involve possible brain damaging effects in a small proportion of pregnancies where the fetus is *Rh*-positive and the mother is *Rh*-negative. (*Rh*-negative has a frequency of 15 percent in white and 7 percent in the Negro population.) The mother-child *Rh* incompatibility produces significant physical ill effects in only a fraction of cases and increases in importance in pregnancies beyond the first. The general finding of slightly lower IQs in second and later born children could be related to *Rh* incompatibility or to similar, but as yet undiscovered, mother-child biological incompatibilities. This is clear an area greatly in need of pioneering research.

Nutrition. Since the human brain attains 70 percent of its maximum adult weight in the first year after birth, it should not be surprising that prenatal and infant nutrition can have significant effects on brain development. Brain growth is largely a process of protein synthesis. During the prenatal period and the first postnatal year the brain normally absorbs large amounts of protein nutrients and grows at the average rate of 1 to 2 milligrams per minute (Stoch & Smythe, 1963; Cravioto, 1966).

Severe undernutrition before two or three years of age, especially a lack of proteins and vitamins and minerals essential for their anabolism, results in lowered intelligence. Stoch and Smythe (1963) found, for example, that extremely malnourished South African colored children were some 20 points lower in IQ than children of similar parents who had not suffered from malnutrition. The difference between the undernourished group and the control group in DQ and IQ over the age range from 1 year to 8 years was practically constant. If undernutrition takes a toll, it takes it early, as shown by the lower DQs at 1 year and the absence of any increase in the decrement at later ages. Undernutrition occurring for the first time in older children seems to have no permanent effect. Severely malnourished war prisoners, for example, function intellectually at their expected level when they are returned to normal living conditions. The study by Stoch and Smythe, like several others (Cravioto, 1966; Scrimshaw, 1968), also revealed that the undernourished children had smaller stature and head circumference than the control children. Although there is no correlation between intelligence and head circumference in normally nourished children, there is a positive correlation between these factors in groups whose numbers suffer varying degrees of undernutrition early in life. Undernutrition also increases the correlation between intelligence and physical stature. These correlations provide us with an index which could aid the study of IQ deficits due to malnutrition in selected populations.

One of the most interesting and pronounced psychological effects of undernutrition is retardation of cross-modal transfer or intersensory integration, which was earlier described as characterizing the essence of *g* (Scrimshaw, 1968).

The earlier the age at which nutritional therapy is instituted, of course, the more beneficial are its effects. But even as late as 2 years of age, a gain of as much as 18 IQ points was produced by nutritional improvements in a group of extremely undernourished children. After 4 years of age, however, nutritional therapy effected no significant change in IQ (Cravioto, 1966, p. 82).

These studies were done in countries where extreme undernutrition is not uncommon. Such gross nutritional deprivation is rare in the United States. But there is at least one study which shows that some undetermined proportion of the urban population in the United States might benefit substantially with respect to intellectual development by improved nutrition. In New York City, women of low socioeconomic status were given vitamin and mineral supplements during pregnancy. These women gave birth th children who, at four years of age, averaged 8 points higher in IQ than a control group of children whose mothers had been given placebos during pregnancy. (Harrell, Woodyard, & Gates, 1955). Vitamin and mineral supplements are, of course, beneficial in this way only when they remedy an existing deficiency.

Birth Order. Order of birth contributes a significant proportion of the variance in mental ability. On the average, first-born children are superior in almost every way, mentally and physically. This is the consistent finding of many studies (Altus, 1966) but as yet the phenomenon remains unexplained. (Rimland [1964, pp. 140-143] has put forth some interesting hypotheses to explain the superiority of the first-born.) Since the first-born effect is found throughout all social classes in many countries and has shown up in studies over the past 80 years (it was first noted by Galton), it is probably a biological rather than a social-psychological phenomenon. It is almost certainly not a genetic effect. (It would tend to make for slightly lower estimates of heritability based on sibling comparisons.) It is one of the sources of environmental variance in ability without any significant postnatal environmental correlates. No way is known for giving later-born children the same advantage. The disadvantage of being later-born, however, is very slight and shows up conspicuously only in the extreme upper tail of the distributions of achievements. For example, there is a disproportionate number of first-born individuals whose biographies appears in *Who's Who* and in the *Encyclopedia Britannica.*

Social Class Differences in Intelligence

Social class (or socioeconomic status [SES]) should be considered as a factor separate from race. I have tried to avoid using the terms *social class* and *race* synonymously or interchangeably in my writings, and I observe this distinction here. Social classes completely cut across all racial groups. But different racial groups are disproportionately represented in different SES categories. Social class differences refer to a socioeconomic continuum *within* racial groups.

It is well known that children's IQs, by school age, are correlated with the socioeconomic status of their parents. This is a worldwide phenomenon and has an extensive research literature going back 70 years. Half of all the correlations between SES and children's IQs reported in the literature fall between .25 and .50, with most falling in the region of .35 to .40. When school children are grouped by SES, the mean IQs of the groups vary over a range of one to two standard deviations (15 to 30 IQ points), depending on the method of status classification (Eells, et al., 1951). This relationship between SES and IQ constitutes one of the most substantial and least disputed facts in psychology and education.

The fact that intelligence is correlated with occupational status can hardly be surprising in any society that supports universal public education. The educational system and occupational hierarchy act as an intellectual "screening" process, far from perfect, to be sure, but discriminating enough to create correlations of the magnitude just reported. If each generation is roughly sorted out by these "screening" processes along an intelligence continuum, and if, as has already been pointed out, the phenotype-genotype correlation for IQ is of the order of .80 to .90, it is almost inevitable that this sorting process with make for genotypic as well as phenotypic differences among social classes. It is therefore most unlikely that groups differing in SES would not also differ, on the average, in their genetic endowment of intelligence. In reviewing the relevant evidence, the British geneticist, C.O. Carter (1966, p. 192) remarked, "Sociologists who doubt this show more ingenuity than judgment." Sociologist Bruce Eckland (1967) has elaborately spelled out the importance of genetic factors for understanding social class differences.

Few if any students of this field today would regard socioeconomic status per se as an environmental variable that primarily *causes* IQ differences. Intellectual differences between SES groups have hereditary, environmental, and interaction components. Environmental factors associated with SES differences apparently are not a major *independent* source of variance in intelligence. Identical twins separated in the first months of life and reared in widely differing social classes, for example, still show greater similarity in IQ than unrelated children reared together or even siblings reared together (Burt, 1966). The IQs of children adopted in infancy show a much lower correlation with the SES of the adopting parents than do the IQs of children reared by their own parents (Leahy, 1935). The IQs of children who were reared in an orphanage from infancy and who had never known their biological parents show approximately the same correlation with their biological father's occupational status as found for children reared by their biological parents (.23 *vs* .24) (Lawrence, 1931). The correlation between the IQs of children adopted in infancy and the educational level of their biological mothers is close to that of children reared by their own mothers (.44), while the correlation between children's IQs and their adopting parents' educational level is close to zero (Honzik, 1957). Children of low and high SES show, on the average, an amount of regression from the parental IQ toward the mean of the general population that conforms to expectations from a simple polygenic model of the inheritance of intelligence (Burt, 1961). When siblings reared within the same family differ significantly in intelligence, those who are above the family average tend to move up the SES scale, and those who are blow the family average tend to move down (Young & Gibson, 1965). It should also be noted that despite intensive efforts by psychologists, educators, and sociologists to devise tests intended to eliminate SES differences in measured intelligence, none of these efforts has succeeded (Jensen, 1968c). Theodosius Dobzhansky (1968a, p. 33), a geneticist, states that "There exist some occupations or functions for which only extreme genotypes are suitable." But surely this is not an all-or-nothing affair, and we would expect by the same reasoning that many different occupational skills, and not just those that are the most extreme, would favor some genotypes more than others. To be sure, genetic factors become more important at the extremes. Some minimal level of ability is required for learning most skills. But while you can teach almost anyone to play chess, or the piano, or to conduct an orchestra, or to write prose, you cannot teach everyone to be a Capablanca, a Pederewski, a Toscanini, or a Bernard Shaw. In a society that values and rewards individual talent and merit, genetic factors inevitably take on considerable importance.

SES differences, and race differences as well, are manifested not only as differences between group means, but also as differences in variance and in patterns of correlations among various mental abilities, even on tests which show no *mean* differences between SES groups (Jensen, 1968b).

Another line of evidence that SES IQ differences are not a superficial phenomenon is the fact of a negative correlation between SES and Developmental Quotient (DQ) (under two years of age) and an increasing positive correlation between SES and IQ (beyond two years of age), as shown in Figure 14 from a study by Nancy Bayley (1966). (All subjects in this study are Caucasian.) This relationship is especially interesting in view of the finding of a number of studies that there is a negative correlation between DQ and later IQ, an effect which is much more pronounced in boys than in girls and involves the motor more than the attentional-cognitive aspects of the DQ (Bayley, 1965b). Figure 14 shows that on infant developmental scales, lower SES children actually have a "head start" over higher SES children. But this trend is increasingly reversed at later ages as the tests become less motoric and are increasingly loaded with a cognitive or *g* factor.

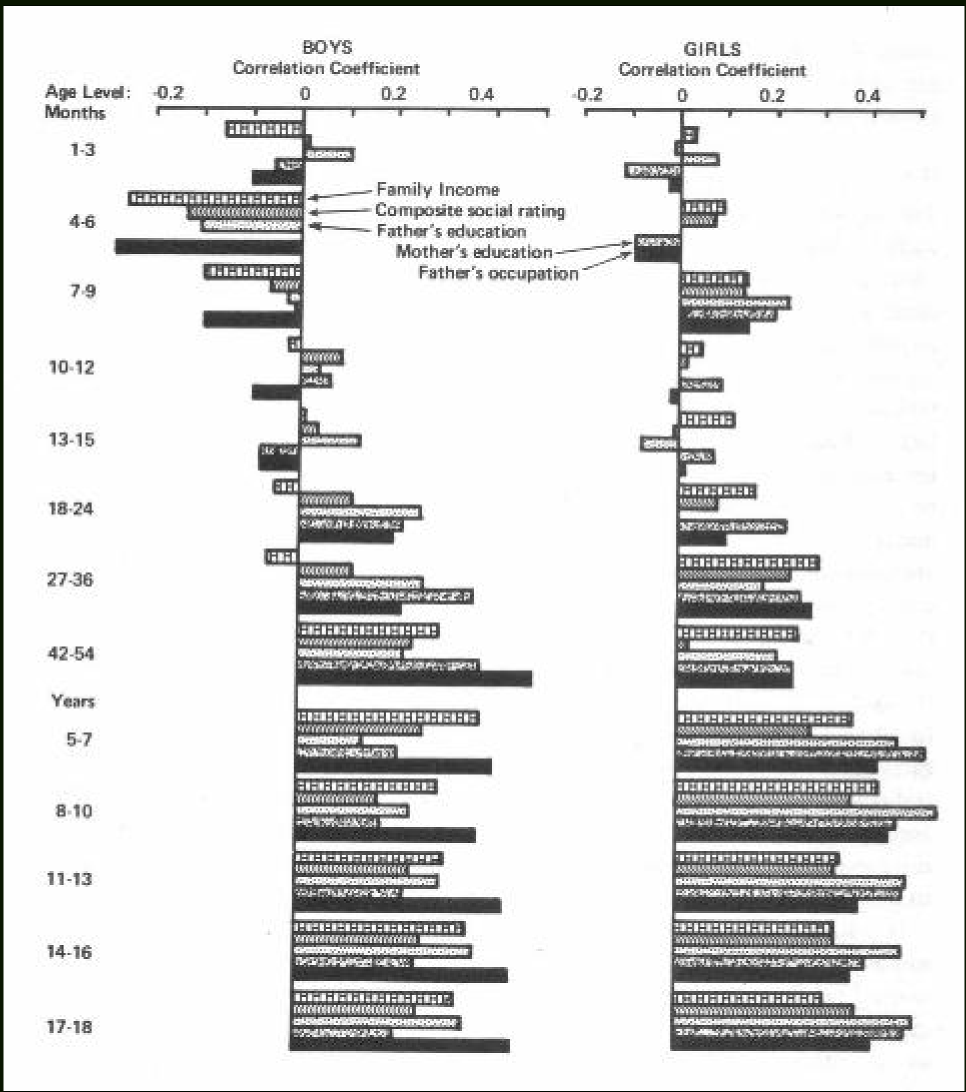

FIGURE 14.
Correlations between children's mental test scores, at 1 month to 18 years, and five indicators of parents' socioeconomic status at the time the children were born. (Bayley, 1966.)

Race Differences

The important distinction between the *individual* and the *population* must always be kept clearly in mind in any discussion of racial differences in mental abilities or any other behavioral characteristics.

Whenever we select a person for some special educational purpose, whether for special instruction in a grade-school class for children with learning problems, or for a "gifted" class with an advanced curriculum, or for college attendance, or for admission to graduate training or a professional school, we are selecting an *individual*, and we are selecting him and dealing with him for reasons of his individuality. Similarly, when we employ someone, or promote someone in his occupation, or give some special award or honor to someone for his accomplishments, we are doing this to an individual. The variables of social class, race, and national origin are correlated so imperfectly with any of the valid criteria on which the above decisions should depend, or, for that matter, with any behavioral characteristic, that these background factors are irrelevant as a basis for dealing with individuals—as students, as employees, as neighbors. Furthermore, since, as far as we know, the full range of human talents is represented in all the major races of man and in all socioeconomic levels, it is unjust to allow the mere fact of an individual's racial or social background to affect the treatment accorded to him. All persons rightfully must be regarded on the basis of their individual qualities and merits, and all social, educational, and economic institutions must have built into them the mechanisms for insuring and maximizing the treatment of persons according to their individual behavior.

If a society completely believed and practiced the ideal of treating every person as an individual, it would be hard to see why there should be any problems about "race" per se. There might still be problems concerning poverty, unemployment, crime, and other social ills, and, given the will, they could be tackled just as any other problems that require rational methods for solution. But if this philosophy prevailed in practice, there would not need to be a "race problem."

The question of *race* differences in intelligence comes up not when we deal with individuals as individuals, but when certain identifiable *groups* or subcultures within the society are brought into comparison with one another *as groups or populations*. It is only when the groups are disproportionately represented in what are commonly perceived as the most desirable and the least desirable social and occupational roles in a society that the question arises concerning average differences among groups. Since much of the current thinking appeals to the fact that there is a disproportionate representation of different racial groups in the various levels of the educational, occupational, and socioeconomic hierarchy, we are forced to examine all the possible reasons for this inequality among racial groups in the attainments and rewards generally valued by all groups within our society. To what extent can such inequalities be attributed to unfairness in society's multiple selection processes? ("Unfair" meaning that selection is influenced by intrinsically irrelevant criteria, such as skin color, racial or national origin, etc.) And to what extent are these inequalities attributable to really relevant selection criteria which apply equally to all individuals but at the same time select disproportionately between some racial groups because there exist, in fact, real average differences among the groups—differences in the population distributions of those characteristics which are indisputably relevant to educational and occupational performance? This certainly is one of the most important questions confronting our nation today. The answer, which can be found only through unfettered research, has enormous consequences for the welfare of all, particularly of minorities whose plight is now in the foreground of public attention. A preordained, doctrinaire stance with regard to this issue hinders the achievement of a scientific understanding of the problem. To rule out of court, so to speak, any reasonable hypotheses on purely ideological grounds is to argue that static ignorance is preferable to increasing our knowledge of reality. I strongly disagree with those who believe in searching for the truth by scientific means only under certain circumstances, or who believe that the results of inquiry on some subjects cannot be entrusted to the public but should be kept the guarded possession of a scientific elite. Such attitudes, in my opinion, represent a danger to free inquiry and, consequently, in the long run, work to the disadvantage of society's general welfare. "No holds barred" is the best formula for scientific inquiry. One does not decree beforehand which phenomena cannot be studied or which questions cannot be answered.

Genetic Aspects of Racial Differences. No one, to my knowledge, questions the role of environmental factors, including influences from past history, in determining at least some of the variance between racial groups in standard measures of intelligence, school performance, and occupational status. The current literature on the culturally disadvantaged abounds with discussion—some of it factual, some of it fanciful—of how a host of environmental factors depresses cognitive development and performance. I recently co-edited a book which is largely concerned with the environmental aspects of disadvantaged minorities

(Deutsch, Katz, & Jensen, 1968). But the possible importance of genetic factors in racial behavioral differences has been greatly ignored, almost to the point of being a tabooed subject, just as were the topics of venereal disease and birth control a generation or so ago.

My discussions with a number of geneticists concerning the question of a genetic basis of differences among races in mental abilities have revealed to me a number of rather consistently agreed-upon points which can be summarized in general terms as follows: Any groups which have been geographically or socially isolated from one another for many generations are practically certain to differ in their gene pools, and consequently are likely to show differences in any phenotypic characteristics having high heritability. This is practically axiomatic, according to the geneticists with whom I have spoken. Races are said to be "breeding populations," which is to say that matings within the group have a much higher probability than matings outside the group. Races are more technically viewed by geneticists as populations having different distributions of gene frequencies. These genetic differences are manifested in virtually every anatomical, physiological, and biochemical comparison one can make between representative samples of identifiable racial groups (Kuttner, 1967). There is no reason to suppose that the brain should be exempt from this generalization. (Racial differences in the relative frequencies of various blood constituents have probably been the most thoroughly studied so far.)

But what about behavior? If it can be measured and shown to have a genetic component, it would be regarded, from a genetic standpoint, as no different from other human characteristics. There seems to be little question that racial differences in genetically conditioned behavioral characteristics, such as mental abilities, should exist, just as physical differences. The real questions, geneticists tell me, are not whether there are or are not genetic racial differences that affect behavior, because there undoubtedly are. The proper question to ask, from a scientific standpoint, are: What is the direction of the difference? What is the magnitude of the difference? And what is the significance of the difference—medically, socially, educationally, or from whatever standpoint that may be relevant to the characteristic in question? A difference is important only within a specified context. For example, one's blood type in the ABO system is unimportant until one needs a transfusion. And some genetic differences are apparently of no importance with respect to any context as far as anyone has been able to discover—for example, differences in the size and shape of ear lobes. The idea that all genetic differences have arisen or persisted only as a result of natural selection, by conferring some survival or adaptive benefit on their possessors, is no longer generally held. There appear to be many genetic differences, or polymorphisms, which confer no discernible advantages to survival.[6]

Negro Intelligence and Scholastic Performance. Negroes in the United States are disproportionately represented among groups identified as culturally or educationally disadvantaged. This, plus the fact that Negroes constitute by far the largest racial minority in the United States, has for many years focused attention on Negro intelligence. It is a subject with a now vast literature which has been quite recently reviewed by Dreger and Miller (1960, 1968) and by Shuey (1966), whose 578 page review is the most comprehensive, covering 382 studies. The basic data are well known: on the average, Negroes test about 1 standard deviation (15 IQ points) below the average of the white population in IQ, and this finding is fairly uniform across the 81 different tests of intellectual ability used in the studies reviewed by Shuey. The magnitude of difference gives a median overlap of 15 percent, meaning that 15 percent of the Negro population exceeds the white average. In terms of proportions of variance, if the numbers of Negroes and whites were equal, the difference *between* racial groups would account for 23 percent of the total variance, but—an important point—the differences *within* groups would account for 77 percent of the total variance. When gross socioeconomic level is controlled, the average differences reduces to about 11 IQ points (Shuey, 1966, p. 519), which, it should be recalled, is about the same spread as the average difference between siblings in the same family. So-called "culture-free" or "culture-fair" tests tend to give Negroes slightly lower scores, on the average, than more conventional IQ tests such as the Stanford-Binet and Wechsler scales. Also, as a group, Negroes perform somewhat more poorly on those subtests which tap abstract abilities. The majority of studies show that Negroes perform relatively better on verbal than on non-verbal intelligence tests.

In tests of scholastic achievement, also, judging from the massive data of the Coleman study (Coleman, et al., 1966), Negroes score about 1 standard deviation (SD) below the average for whites and Orientals and considerably less than 1 SD below other disadvantaged minorities in the Coleman study—Puerto Rican, Mexican-American, and American Indian. The 1 SD decrement in Negro performance is fairly constant throughout the period from grades 1 through 12.

Another aspect of the distribution of IQs in the Negro population is their lesser variance in comparison to the white distribution. This shows up in most of the studies reviewed by Shuey. The best single estimate is probably the estimate based on a large normative study of Stanford-Binet IQs of Negro school children in five Southeastern states, By Kennedy, Van De Riet, and White (1963). They found the SD of Negro children's IQs to be 12.4, as compared with 16.4 in the white normative sample. The Negro distribution thus has only about 60 percent as much variance (i.e., SD^2) as the white distribution.

There is an increasing realization among students of the psychology of the disadvantaged that the discrepancy in their average performance cannot be completely or directly attributed to discrimination or inequalities in education. It seems not unreasonable, in view of the fact that intelligence variation has a large genetic component, to hypothesize that genetic factors may play a part in this picture. But such an hypothesis is anathema to many social scientists. The idea that the lower average intelligence and scholastic performance of Negroes could involve, not only environmental, but also genetic, factors has indeed been strongly denounced (e.g., Pettigrew, 1964). But it has been neither contradicted nor discredited by evidence.

The fact that a reasonable hypothesis has not been rigorously proved does not mean that it should be summarily dismissed. It only means that we need more appropriate research for putting it to the test. I believe such definitive research is entirely possible but has not yet been done. So all we are left with are various lines of evidence, no one of which is definitive alone, but which, viewed all together, make it a not unreasonable hypothesis that genetic factors are strongly implicated in the average Negro-white intelligence difference. The preponderance of the evidence is, in my opinion, less consistent with a strictly environmental hypothesis than with a genetic hypothesis, which, of course, does not exclude the influence of environment or its interaction with genetic factors.

We can be accused of superficiality in our thinking about this issue, I believe, if we simply dismiss a genetic hypothesis without having seriously thought about the relevance of typical findings such as the following.

Failure to Equate Negroes and Whites in IQ and Scholastic Ability. No one has yet produced any evidence based on a properly controlled study to show that representative samples of Negro and white children can be equalized in intellectual ability through statistical control of environment and education.

Socioeconomic Level and Incidence of Mental Retardation. Since in no category of socioeconomic status (SES) are a majority of children found to be retarded in the technical sense of having an IQ below 75, it would be hard to claim that the degree of environmental deprivation typically associated with lower-class status could be responsible for this degree of mental retardation. An IQ less than 75 reflects more than a lack of cultural amenities. Heber (1968) has estimated on the basis of existing evidence that IQs below 75 have a much hither incidence among Negro than among white children at every level of socioeconomic status, as shown in Table 3. In the two highest SES categories the estimated proportions of Negro and white children with IQs below 75, are in the ratio of 13.6 to 1. If environmental factors were mainly responsible for producing such differences, one should expect a lesser Negro-white discrepancy at the upper SES levels. Other lines of evidence also show this not to be the case. A genetic hypothesis, on the other hand, would predict this effect, since the higher SES Negro offspring would be regressing to a lower population mean than their white counterparts in SES, and consequently a larger proportion of the lower tail of the distribution of genotypes for Negroes would fall below the value that generally results in phenotypic IQs below 75.

TABLE 3
Estimated Prevalence of Children With IQs Below 75, by Socioeconomic Status (SES) and

Race Given as Percentages (Heber, 1968)

SES	*White*	*Negro*
High 1	0.5	3.1
2	0.8	14.5
3	2.1	22.8
4	3.1	37.8
Low 5	7.8	42.9

A finding reported by Wilson (1967) is also in line with this prediction. He obtained the mean IQs of a large representative sample of Negro and white children in a California school district and compared the two groups within each of four social class categories: (1) professional and managerial, (2) white collar, (3) skilled and semiskilled manual, and (4) lower class (unskilled, unemployed, or welfare recipients). The mean IQ of Negro children in the first category was 15.5 points below that of the corresponding white children in SES category 1. But the Negro mean for SES 1 was also 5.9 points below the mean of white children in SES category 4. (The IQs of white children in SES 4 presumably have "regressed" upward toward the mean of the white population.)

Wilson's data are not atypical, for they agree with Shuey's (1966, p. 520) summarization of the total literature up to 1965 on this point. She reports that in all the studies which grouped subjects by SES, upper-status Negro children average 2.6 IQ points *below* the low status whites. Shuey comments: "It seems improbable that upper and middle-class colored children would have no more culture opportunities provided them than white children of the lower and lowest class."

Duncan (1968, p. 69) also has presented striking evidence for a much greater "regression-to-the-mean" (from parents to their children) for high status occupations in the case of Negroes than in the case of whites. None of these findings is at all surprising from the standpoint of a genetic hypothesis, of which an intrinsic feature is Galton's "law of filial regression." While the data are not necessarily inconsistent with a possible environmental interpretation, they do seem more puzzling in terms of a strictly environmental causation. Such explanations often seem intemperately strained.

Inadequacies of Purely Environmental Explanations. Strictly environmental explanations of group differences tend to have an ad hoc quality. They are usually plausible for the situation they are devised to explain, but often they have little generality across situations, and new ad hoc hypotheses have to be continually devised. Pointing to environmental differences between groups is never sufficient in itself to infer a causal relationship to group differences in intelligence. To take just one example of this tendency of social scientists to attribute lower intelligence and scholastic ability to almost any environmental difference that seems handy, we can look at the evidence regarding the effects of "father absence." Since the father is absent in a significantly larger proportion of Negro than of white families, the factor of "father absence" has been frequently pointed to in the literature on the disadvantaged as one of the causes of Negroes' lower performance on IQ tests and in scholastic achievement. Yet the two largest studies directed at obtaining evidence on this very point—the only studies I have seen that are methodologically adequate—both conclude that the factor of "father absence" versus "father presence" makes no independent contribution to variance in intelligence or scholastic achievement. The sample sizes were so large in both of these studies that even a very slight degree of correlation between father-absence and the measures of cognitive performance would have shown up as statistically significant. Coleman (1966, p. 506) concluded: "Absence of a father in the home did not have the anticipated effect on ability scores. Overall, pupils without fathers performed at approximately the same level as those with fathers—although there was some variation between groups" (groups referring to geographical regions of the U.S.). And Wilson (1957, p. 177) concluded from his survey of a California school district: "Neither our own data nor the preponderance of evidence from other research studies indicate that father presence or absence *per se*, is related to school achievement. While broken homes reflect the existence of social and personal problems, and have some consequence for the development of personality, broken homes do not have any systematic effect on the overall level of school success."

The nationwide Coleman study (1966) included assessments of a dozen environmental variables and socioeconomic indices which are generally thought to be major sources of environmental influence in determining individual and group differences in scholastic performance—such factors as: reading material in the home, cultural amenities in the home, structural integrity of the home, foreign language in the home, preschool attendance, parents' education, parents' educational desires for child, parents' interest in child's school work, time spent on homework, child's self-concept (self-esteem), and so on. These factors are all correlated—in the expected direction—with scholastic performance within each of the racial or ethnic groups studied by Coleman. Yet, interestingly enough, they are not systematically correlated with differences *between* groups. For example, by far the most economically disadvantaged groups in the Coleman study are the American Indians. On every environmental index they average *lower* than the Negro samples, and overall their environmental rating is about as far below the Negro average as the Negro rating is below the White average. (As pointed out by Kuttner [1968, p. 707], American Indians are much more disadvantaged than Negroes, or any other minority groups in the United States, on a host of other factors not assessed by Coleman, such as income, unemployment, standards of health care, life expectancy, and infant mortality.) Yet the American Indian ability and achievement test scores average about half a standard deviation higher than the scores of Negroes. The differences were in favor of the Indian children on each of the four tests used by Coleman: non-verbal intelligence, verbal intelligence, reading comprehension, and math achievement. If the environmental factors assessed by Coleman are the major determinants of Negro-white differences that many social scientists have claimed they are, it is hard to see why such factors should act in reverse fashion in determining differences between Negroes and Indians, especially in view of the fact that *within* each group the factors are significantly correlated in the expected direction with achievement.

Early Developmental Differences. A number of students of child development have noted the developmental precocity of Negro infants, particularly in motoric behavior. Geber (1958) and Geber and Dean (1957) have reported this precocity also in African infants. It hardly appears to be environmental, since it is evident in nine-hour-old infants. Cravioto (1966, p. 78) has noted that the Gesell tests of infant behavioral development, which are usually considered suitable only for children over four weeks of age, "can be used with younger African, Mexican, and Guatemalan infants, since their development at two or three weeks is similar to that of Western European infants two or three times as old." Bayley's (1965a) study of a representative sample of 600 American Negro infants up to 15 months of age, using the Bayley Infant Scales of Mental and Motor Development, also found Negro infants to have significantly higher scores than white infants in their first year. The difference is largely attributable to the motor items in the Bayley test. For example, about 30 percent of white infants as compared with about 60 percent of Negro infants between 9 and 12 months were able to "pass" such tests as "pat-a-cake" muscular coordination, and ability to walk with help, to stand alone, and to walk alone. The highest scores for any group on the Bayley scales that I have found in my search of the literature were obtained by Negro infants in the poorest sections of Durham, North Carolina. The older siblings of these infants have an average IQ of about 80. The infants up to 6 months of age, however, have a Developmental Motor Quotient (DMQ) nearly one standard deviation above white norms and a Developmental IQ (i.e., the non-motor items of the Bayley scale) of about half a standard deviation above white norms (Durham Education Improvement Program, 1966-67, a, b).

The DMQ, as pointed out previously, correlates negatively in the white population with socioeconomic status and with later IQ. Since SES Negro and white school children are more alike in IQ than are upper SES children of the two groups (Wilson, 1967), one might expect greater DMQ differences in favor of Negro infants in high socioeconomic Negro and white samples than in low socioeconomic samples. This is just what Walters (1967) found. High SES Negro infants significantly exceeded whites in total score on the Gesell developmental schedules at 12 weeks of age, while low SES Negro and white infants did not differ significantly overall. (The only difference, on a single subscale, favored the white infants.)

It should also be noted that developmental quotients are usually depressed by adverse prenatal, perinatal, and postnatal complications such as lack of oxygen, prematurity, and nutritional deficiency.

Another relationship of interest is the finding that the negative correlation between DMQ and later IQ is higher in boys than in girls (Bayley, 1966, p. 127). Bronfenbrenner (1967, p. 912) cites evidence which shows that Negro boys perform relatively less well in school than Negro girls; the sex difference is much greater than is found in the white population. Bronfenbrenner (1967, p. 913) says, "It is noteworthy that these sex differences in achievement are observed among Southern as well as Northern Negroes, are present at every socioeconomic level, and tend to increase with age."

Physiological Indices. The behavioral precocity of Negro infants is also paralleled by certain physiological indices of development. For example, x-rays show that bone development, as indicated by the rate of ossification of cartilage, is more advanced in Negro as compared with white babies of about the same socioeconomic background, and Negro babies mature at a lower birth-weight than white babies (Naylor & Myrianthopoulos, 1967, p. 81).

It has also been noted that brain wave patterns in African newborn infants show greater maturity than is usually found in the European newborn child (Nilson & Dean, 1959). This finding especially merits further study, since there is evidence that brain waves have some relationship to IQ (Medical World News, 1968), and since at least one aspect of brain waves—the visually evoked potentially—has a very significant genetic component, showing a heritability of about .80 (uncorrected for attenuation) (Dustman & Beck, 1965).

Magnitude of Adult Negro-White Differences. The largest sampling of Negro and white intelligence test scores resulted from the administration of the Armed Forces Qualification Test (AFQT) to a national sample of over 10 million men between the ages of 18 and 26. As of 1966, the overall failure rate for Negroes was 68 percent as compared with 19 percent for whites (*U.S. News and World Report,* 1966). (The failure cut-off score that yields these percentages is roughly equivalent to a Stanford-Binet IQ of 86.) Moynihan (1965) has estimated that during the same period in which AFQT was administered to these large representative samples of Negro and white male youths, approximately one-half of Negro families could be considered as middle-class or above by the usual socioeconomic criteria. So even if we assumed that all of the lower 50 percent of Negroes on the SES scale failed the AFQT, it wold still mean that at least 36 percent of the middle SES Negroes failed the test, a failure rate almost twice as high as that of the white population at all levels of SES.

Do such findings raise any question as to the plausibility of theories that postulate exclusively environmental factors as sufficient causes for the observed differences?

Why Raise Intelligence?

If the intelligence of the whole population increased and our IQ tests were standardized anew, the mean IQ would again be made equal to 100, which, by definition, is the average for the population. Thus, in order to speak sensibly of raising intelligence we need an absolute frame of reference, and for simplicity's sake we will use the *present* distribution of IQ as our reference scale. Then it will not be meaningless to speak of the average IQ of the population shifting to values other than 100.

Would there be any real advantage to shifting the entire distribution of intelligence upward? One way to answer this question is to compare the educational attainments of children in different schools whose IQ distributions center around means of, say, 85, 100, and 115. As pointed out earlier, there is a relationship between educational attainments and the occupations that are open to individuals on leaving school. Perusal of the want-ads in any metropolitan newspaper reveals that there are extremely few jobs advertised which are suitable to the level of education and skills typically found below IQs of 85 or 90, while we see day after day in the want-ads hundreds of jobs which call for a level of education and skills typically found among school graduates with IQs above 110. These jobs go begging to be filled. The fact is, there are not nearly enough minimally qualified persons to fill them.

One may sensibly ask the question whether our collective national intelligence is adequate to meet the growing needs of our increasingly complex industrial society. In a bygone era, when the entire population's work consisted almost completely of gathering or producing food by primitive means, there was little need for a large number of persons with IQs much above 100. Few of the jobs that had to be done at that time required the kinds of abstract intelligence and academic training which are now in such seemingly short supply in relation to the demand in our modern society. For many years the criterion for mental retardation was an IQ below 75. In recent years the National Association for Mental Retardation has raised the criterion to an IQ of 85, since an increasing proportion of persons of more than 1 standard deviation below the average in IQ are unable to get along occupationally in today's world. Persons with IQs of 85 or less are finding it increasingly difficult to get jobs, any jobs, because they are unprepared, for whatever reason, to do the jobs that need doing in this industrialized, technological economy. Unless drastic changes occur—in the population, in educational outcomes, or in the whole system of occupational training and selection—it is hard to see how we can avoid an increase in the rate of the so-called "hard-core" unemployed. It takes more knowledge and cleverness to operate, maintain, or repair a tractor than to till a field by hand, and it takes more skill to write computer programs than to operate an adding machine. And apparently the trend will continue.

It has been argued by Harry and Margaret Harlow that "human beings in our world today have no more, or little more, than the absolute minimal intellectual endowment necessary for achieving the civilization we know today" (Harlow & Harlow, 1962, p. 34). They depict where we would probably be if man's average genetic endowment for intelligence had never risen above the level corresponding to IQ 75: "...the geniuses would barely exceed our normal or average level; comparatively few would be equivalent in ability to our average high school graduates. There would be no individuals with the normal intellectual capacities essential for making major discoveries, and there could be no civilization as we know it."

It may well be true that the kind of ability we now call intelligence was needed in a certain percentage of the human population for our civilization to have arisen. But while a small minority—perhaps only one or two percent—of highly gifted individuals were needed to advance civilization, the vast majority were able to assimilate the consequences of these advances. It may take a Leibnitz or a Newton to invent the calculus, but almost any college student can learn it and use it.

Since intelligence (meaning *g*) is not the whole of human abilities, there may be some fallacy and some danger in making it the *sine qua non* of fitness to play a productive role in modern society. We should not assume certain ability requirements for a job without establishing these requirements as a fact. How often do employment tests, Civil Service examinations, the requirements of a high school diploma, and the like, constitute hurdles that are irrelevant to actually performing on the job for which they are intended as a screening device? Before going overboard in deploring the fact that disadvantaged minority groups fail to clear many of the hurdles that are set up for certain jobs, we should determine whether the educational and mental test barriers that stand at the entrance to many of these employment opportunities are actually relevant. They may be relevant only in the correlational sense that the test predicts success on the job, in which case we should also know whether the test measures the ability actually required on the job or measures only characteristics that happen to be correlated with some third factor which is really essential for job performance. Changing people in terms of the really essential requirements of a given job may be much more feasible than trying to increase their abstract intelligence or level of performance in academic subjects so that they can pass irrelevant tests.

IQ Gains from Environmental Improvement

As was pointed out earlier, since the environment acts as a threshold variable with respect to IQ, an overall increase in IQ in a population in which a great majority are above the threshold, such that most of the IQ variance is due to heredity, could not be expected to be very large if it had to depend solely upon improving the environment of the economically disadvantaged. This is not to say that such improvement is not to be desired for its own sake or that it would not boost the educational potential of many disadvantaged children. An unrealistically high upper limit of what one could expect can be estimated from figures given by Schwebel (1968, p. 210). He estimates that 26 percent of the children in the population can be called environmentally deprived. He estimates the frequencies of their IQs in each portion of the IQ scale; their distribution is skewed, with higher frequencies in the lower IQ categories and an overall mean IQ of 90. Next, he assumes we could add 20 points to each deprived child's IQ by giving him an abundant environment. (The figure of 20 IQ points comes from Bloom's [1964, p. 89] estimate that the effect of extreme environments on intelligence is about 20 IQ points.) The net effect of this 20-point boost in the IQ of every deprived child would be an increase in the population's IQ from 100 to 105. but this seems to be an unrealistic fantasy. For if it were true that the IQs of the deprived group could be raised 20 points by a good environment, and if Schwebel's estimate of 26 percent correctly represents the incidence of deprivation, then the deprived children would be boosted to an average IQ of 110, which is 7 points higher than the mean of 103 for the non-deprived population! There is no reason to believe that the IQs of deprived children, given an environment of abundance, would rise to a higher level than the already privileged children's IQs. The overall boost in the population IQ would probably be more like 1 or 2 IQ points rather than 5. (Another anomaly of Schwebel's "analysis" is that after a 20-point IQ boost is granted to the deprived segment of the population, the only persons left in the mentally retarded range are the non-deprived, with 7 percent of them below IQ 80 as compared with zero percent of the deprived!)

Fewer persons, however, are seriously concerned about whether or not we could appreciably boost the IQ of the population as a whole. A more feasible and urgent goal is to foster the educational and occupational potential of the disadvantaged segment of the population. The pursuit of this aim, of course, must involve advances not only in education, but in public health, in social services, and in welfare and employment practices. In considering all feasible measures, one must also take inventory of forces that may be working against the accomplishment of amelioration. We should not overlook the fact that social and economic conditions not only have direct environmental effects, but indirectly can have biological consequences as well, consequences that could oppose attempts to improve the chances of the disadvantaged to assume productive roles in society.

Possible Dysgenic Trends

In one large Midwestern city it was found that one-third of all the children in classes for the mentally retarded (IQ less than 75) came from one small area of the city comprising only five percent of the city's population (Heber, 1968). A representative sample of 88 mothers having at least one school-age child in this neighborhood showed an average of 7.6 children per mother. In families of 8 or more, nearly half the children over 12 years of age had IQs below 75 (Heber, Dener, & Conry, 1958). The authors note that not all low SES families contributed equally to the rate of mental retardation in this area; certain specifiable families had a greatly disproportionate number of retarded children. Mothers with IQs below 80, for example, accounted for over 80 percent of the children with IQs under 80. Completely aside from the hereditary implications, what does this mean in view of studies of foster children which show that the single most important factor in the child's *environment* with respect to his intellectual development is his foster mother's IQ? This variable has been shown to make the largest *independent* contribution to variance in children's IQs of any environmental factor (Burks, 1928). If the children in the neighborhoods studied by Heber, which are typical of the situation in many of our large cities, have the great disadvantage of deprived environments, is it inappropriate to ask the same question that Florence Goodenough (1940, p. 329) posed regarding causal factors in retarded Tennessee mountain children: "*Why* are they so deprived?" When a substantial proportion of the children in a community suffer a deplorable environment, one of the questions we need to answer is who creates their environment? Does not the genetic x environment interaction work both ways, the genotype to some extent making its own environment and that of its progeny?

In reviewing evidence from foster home studies on environmental amelioration of IQs below 75 (the range often designated as indicating cultural-familial retardation) Heber, Dever, and Conry (1968, p. 17) state: "The conclusion that changes in the living environment can cause very large increments in IQ *for the cultural-familial retardate* is not warranted by these data."

What is probably the largest study ever made of familial influences in mental retardation (defined in this study as IQ less than 70) involved investigation of more than 80,000 relatives of a group of mentally retarded persons by the Dight Institute of Genetics, University of Minnesota (Reed & Reed, 1965). From this large-scale study, Sheldon and Elizabeth Reed estimated that about 80 percent of mentally retarded (IQ less than 70) persons in the United States have a retarded parent or a normal parent who has a retarded sibling. The Reeds state: "One inescapable conclusion is that the transmission of mental retardation from parent to child is by far the most important *single* factor in the persistence of this social misfortune" (p. 48). "The transmission of mental retardation from one generation to the next, should, therefore, receive much more critical attention than it has in the past. It seems fair to state that this problem has been largely ignored on the assumption that if our social agencies function better, that if everyone's environment were improved sufficiently, then mental retardation would cease to be a major problem" (p. 77).

An interesting sidelight of the Reeds' study is the finding that in a number of families in which one or both parents had IQs below 70 and in which the environment they provided their children was deplorably deprived, there were a few children of average and superior IQ (as high as 130 or above) and superior scholastic performance. From a genetic standpoint the occurrence of such children would be expected. It is surprising from a strictly environmental standpoint. But, even though some proportion of the children of retarded parents are obviously intellectually well endowed, who would wish upon them the kind of environment typically provided by retarded parents? An investigation conducted in Denmark concluded that "...it is a very severe psychical trauma for a normally gifted child to grow up in a home where the mother is mentally deficient" (Jepsen & Bredmose, 1956, p. 209). Have we thought sufficiently of the rights of children—of their right to be born with fair odds against being mentally retarded, not to have a retarded parent, and with fair odds in favor of having the genetic endowment needed to compete on equal terms with the majority of persons in society. Can we reasonably and humanely oppose such rights of millions of children as yet not born?

Is Our National IQ Declining? It has long been known that there is a substantial negative correlation (averaging about -.30 in various studies) between intelligence and family size and between social class and family size (Anastasi, 1956). Children with many siblings, on the average have lower IQs than children in small families, and the trend is especially marked for families of more than five (Gottesman, 1968). This fact once caused concern in the United States, and even more so in Britain, because of its apparent implication of a declining IQ in the population. If more children are born to persons in the lower half of the intelligence distribution, one would correctly predict a decline in the average IQ of the population. In a number of large-scale studies addressed to the issue in Britain and the United States some 20 years ago, no evidence was found for a general decline in IQ (Duncan, 1952). The paradox of the apparent failure of the genetic prediction to be manifested was resolved to the satisfaction of most geneticists by three now famous studies, one by Higgins, Reed, and Reed (1962), the others by Bajema (1963, 1966). All previous analyses had been based on IQ comparisons of children having different numbers of siblings, and this was their weakness. The data needed to answer the question properly consisted of the average number of children born to *all* individuals at every level of IQ. It was found that in the three studies that if persons with very low IQs married and had children, they typically had a large number of children. *But*—it was also found that relatively few persons in the lower tail of the IQ distribution ever married or produced children, and so their reproduction rate is more than counterbalanced by persons at the upper end of the IQ scale, nearly all of whom marry and have children. The data of these studies are shown in Figure 15.

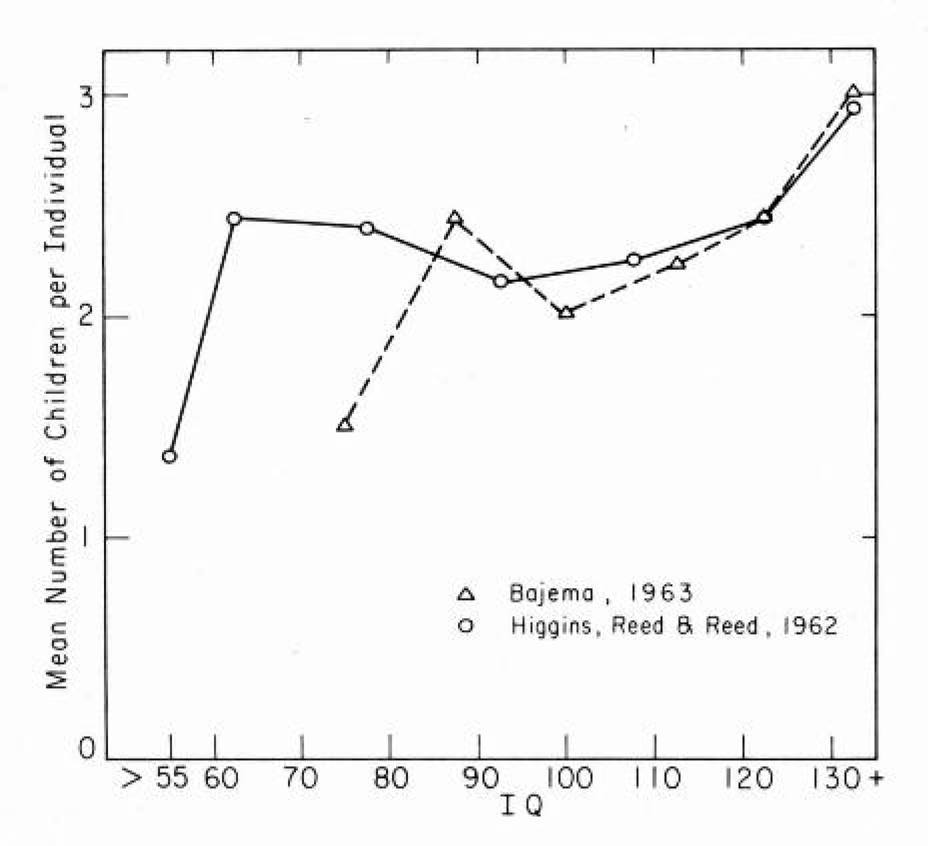

FIGURE 15.
Mean number of children per adult individual (including those who are childless) at each level of IQ, in two samples of white American populations. Note in each sample the bimodal relationship between fertility and IQ.

In my opinion these studies are far from adequate to settle this issue and thus do not justify complacency. They cannot be generalized much beyond the particular generation which the data represent or to other than the white population on which these studies were based. The population sampled by Bajema (1963, 1966), for example, consisted of native-born American whites, predominantly Protestant, with above-average educational attainments, living all or most of their lives in an urban environment, and having most of their children before World War II. Results from a study of this population cannot be confidently generalized to other, quite dissimilar segments of our national population. The relationship between reproductive rate and IQ found by Bajema and by Higgins et al. may very well not prevail in every population group. Thus the evidence to date has not nullified the question of whether dysgenic trends are operating in some sectors.

If this conclusion is not unwarranted, then our lack of highly relevant information on this issue with respect to our Negro population is deplorable, and no one should be more concerned about it than the Negro community itself. Certain census statistics suggest that there might be forces at work which could create and widen the genetic aspect of the average difference in ability between the Negro and white populations in the United States, with the possible consequence that the improvement of educational facilities and increasing equality of opportunity will have a *decreasing* probability of producing equal achievement or continuing gains in the Negro population's ability to compete on equal terms. The relevant statistics have been presented by Moynihan (1966). The differential birthrate, as a function of socioeconomic status, is greater in the Negro than in the white population. The data showing this relationship for one representative age group from the U.S. Census of 1960 are presented in Figure 16.

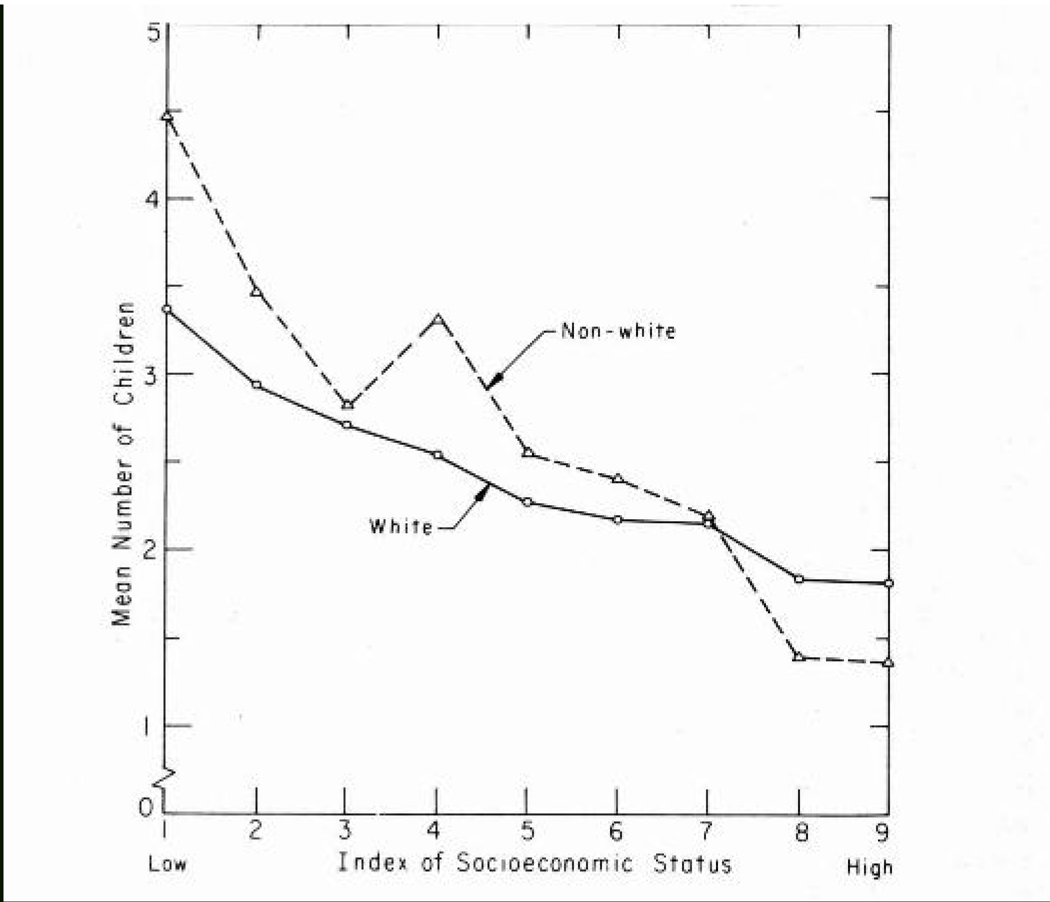

FIGURE 16.
Average number of children per woman 25 to 29 years of age, married once, with husband present, by race and socioeconomic status. From 1960 U.S. Census. (After Mitra, 1966.)

Negro middle- and upper-class families have fewer children than their white counterparts, while Negro lower-class families have more. In 1960, Negro women of ages 25 to 44 married to unskilled laborers had 4.7 children as compared with 3.8 for non-Negro women in the same situation. Negro women married to professional or technical workers had only 1.9 children as compared with 2.4 for white women in the same circumstances. Negro women with annual incomes below $2000 averaged 5.3 children. The poverty rate for families with 5 or 6 children is 3 1/2 times as high as that for families with one or two children (Hill & Jaffe, 1966). That these figures have some relationship to intellectual ability is seen in the fact that 3 out of 4 Negroes failing the Armed Forces Qualification Test come from families of four or more children.

Another factor to be considered is average generation time, defined as the number of years it takes for the parent generation to reproduce its own number. This period is significantly less in the Negro than in the white population. Also, as noted in the study of Bajema (1966), generation length is inversely related to educational attainment and occupational status; therefore a group with shorter generation length is more likely subject to a possible dysgenic effect.

Much more thought and research should be given to the educational and social implications of these trends for the future. Is there a danger that current welfare policies, unaided by eugenic foresight, could lead to the genetic enslavement of a substantial segment of our population? The possible consequences of our failure seriously to study these questions may well be viewed by future generations as our society's greatest injustice to Negro Americans.

Intensive Educational Intervention

We began with mention of several of the major compensatory education programs and their general lack of success in boosting the scholastic performance of disadvantaged children. It has been claimed that such mammoth programs have not been adequately pinpointed to meeting specific, fine-grained cultural and cognitive needs of these children and therefore should not be expected to produce the gains that could result from more intensive and more carefully focused programs in which maximum cultural enrichment and instructional ingenuity are lavished on a small group of children by a team of experts.

The scanty evidence available seem to bear this out. While massive compensatory programs have produced no appreciable gains in intelligence or achievement (as noted on pp. 2-3), the majority of small-scale experiments in boosting the IQ and educational performance of disadvantaged children have produced significant gains. It is interesting that the magnitude of claimed gains generally decreases as one proceeds from reports in the popular press, to informal verbal reports heard on visits at research sites and in private correspondence, to papers read at meetings, to published papers without presentation of supporting data, and to published papers with supporting data. I will confine my review to some of the major studies in the last category.

First, some general observations.

Magnitude of the gains. The magnitude of IQ and scholastic achievement gains resulting from enrichment and cognitive stimulation programs authentically range between 5 and 20 points for IQs, and between about one-half to two standard deviations for specific achievement measures (reading, arithmetic, spelling, etc.). Heber (1968) reviewed 29 intensive preschool programs for disadvantaged children and found they resulted in an average gain in IQ (at the time of children's leaving the preschool program) of between 5 to 10 points; the average gain was about the same for children whose initial IQs were below 90 as for those of 90 and above.

The amount of gain is related to several factors. The intensity and specificity of the instructional aspects of the program seem to make a difference. Ordinary nursery school attendance, with a rather diffuse enrichment program but with little effort directed at development of specific cognitive skills, generally results in a gain of 5 or 6 IQ points in typical disadvantaged preschoolers. If special cognitive training, especially in verbal skills, is added to the program, the average gain is about 10 points—slightly more or less depending on the amount of verbal content in the tests. Average gains rarely go above this, but when the program is extended beyond the classroom into the child's home, and there is intensive instruction in specific skills under short but highly attention-demanding daily sessions, as in the Bereiter-Engelmann program (1966), about a third of the children have shown gains of as much as 20 points.

Average gains of more than 10 of 15 points have not be obtained on any sizeable groups or been shown to persist or to be replicable in similar groups, although there have been claims that average gains of 20 or more points can be achieved by removing certain cultural and attitudinal barriers to learning. The actual evidence, however, warrants the caution expressed by Bereiter and Engelmann (1966, p. 7): "'Miracle cures' of this kind are sometimes claimed to work with disadvantaged children, as when a child is found to gain 20 points or so in IQ after a few months of preschool experience. Such enormous gains, however, are highly suspect to anyone who is familiar with mental measurements. It is a fair guess that the child could have done as well on the first test except that he misinterpreted the situation, was frightened or agitated, or was not used to responding to instructions. Where genuine learning is concerned, enormous leaps simply do not occur, and leaps of any kind do not occur without sufficient cause."

The initial IQ on entering also has some effect, and this fact may be obscured if various studies are coarsely grouped. Bereiter and Engelmann (1966, p. 16), in analyzing results from eight different preschools for culturally disadvantaged children that followed traditional nursery school methods, concluded that the children's average gain in IQ is *half* the way from their initial IQ level to the normal level of 100. This rule was never more than 2 points in error for the studies reviewed. This same amount of IQ gain is generally noted in disadvantaged children during their first year in regular kindergarten (Brison, 1967, p. 8).

I have found no evidence of comparable gains in non-disadvantaged children. Probably the exceedingly meager gains in some apparently excellent preschool programs for the "disadvantaged" are attributable to the fact that the children in them did not come from a sufficiently deprived home background. Such can be the case when the children are admitted to the program on the basis of "self-selection" by their parents. Parents who seek out a nursery school of volunteer their children for an experimental preschool are more apt to have provided their children with a somewhat better environment than would be typical for a randomly selected group of disadvantaged children. This seems to have been the case in Martin Deutsch's intensive preschool enrichment program at the Institute of Developmental Studies in New York (Powledge, 1967). Both the experimental group (E) and the self-selected control groups (C_{88}) were made up of Negro children from a poor neighborhood in New York City whose parents applied for their admission to the program. The E group received intensive educational attention in what is overall the most comprehensive and elaborate enrichment program I know of. The C_{88} group, of course, received no enriched education. The initial average Stanford-Binet IQs of the E and C_{88} groups were 93.32 and 94.69, respectively. After two years in the enrichment program, the E group had a mean IQ of 95.53 and the C_{88} group had 96.52. Both pre- and post-test differences are nonsignificant. The enrichment program continued for a third year through the first grade. For the children in the E group who had had three years of enrichment, there was a significant gain over the C group of 8 months in reading achievement by the end of the first grade, a score above national norms. This result is in keeping with the general finding that enrichment shows a greater effect on scholastic achievement than on IQ per se.

Many studies have employed no control group selected on exactly the same basis as the experimental group. This makes it virtually impossible to evaluate the effect of the treatment on pre-test—post-test gain, and the problem is made more acute by the fact that enrichment studies often pick their subjects on the basis of their being below the average IQ of the population of disadvantaged children from which they are selected. This makes statistical regression a certainty—the group's mean will increase by an appreciable amount because of the imperfect correlation between test-retest scores over, say, a one-year interval. Since this correlation is known to be considerably lower in younger children than in older children, there will be a considerably greater "gain" due to regression for younger groups of children. The net results of selecting especially backward children on the basis of IQ is that a gain in IQ can be predicted which is not at all attributable to the educational treatment given to the children. Studies using control groups nearly always show this gain in the control group, and only by subtracting the control group's gain from the experimental group's gain can we evaluate the magnitude of the treatment effect. Only the gain over and above that attributable to regression really counts.

Still another factor is involved in the inverse relationship generally found between children's age and the size of IQ gains in an enrichment program. Each single item gotten right in a test like the Stanford-Binet adds increasingly smaller increments to the IQ as children get older. Each Stanford-Binet test item, for example, is worth two months of mental age. At four years of age getting just two additional items right will boost an IQ of 85 up to 93. The same absolute amount of improvement in test performance at 10 years of age would boost an IQ of 85 up to only 88. The typical range of gains found in preschool enrichment programs, in the age range of 4 to 6, are about what would be expected from passing an additional two to four items in the Stanford-Binet. This amount of gain should not be surprising on a test which, for this age range, consists of items rather similar to the materials and activities traditionally found in nursery schools—blocks, animal pictures, puzzles, bead stringing, copying drawings, and the like. I once visited an experimental preschool using the Stanford-Binet to assess pre-test—post-test gains, in which some of the Stanford-Binet test materials were openly accessible to the children throughout their time in the school as part of the enrichment paraphernalia. Years ago Reymert and Hinton (1940) noted this "easy gain" in the IQs of culturally disadvantaged preschoolers on the tests depending on specific information such as being able to name parts of the body and knowing names of familiar objects. Children who have not picked up this information at home get it quickly in nursery school and kindergarten.

In addition to these factors, something else operates to boost scores five to ten points from first to second test, provided the first test is really the first. When I worked in a psychological clinic, I had to give individual intelligence tests to a variety of children, a good many of whom came from an impoverished background. Usually I felt these children were really brighter than their IQ would indicate. They often appeared inhibited in their responsiveness in the testing situation on their first visit to my office, and when this was the case I usually had them come in on two to four different days for half-hour sessions with me in a "play therapy" room, in which we did nothing more than get better acquainted by playing ball, using finger paints, drawing on the blackboard, making things out of clay, and so forth. As soon as the child seemed to be completely at home in this setting, I would retest him on a parallel form of the Stanford-Binet. A boost in IQ of 8 to 10 points or so was the rule; it rarely failed, but neither was the gain very often much above this. So I am inclined to doubt that IQ gains up to this amount in young disadvantaged children have much of anything to do with changes in ability. They are largely a result simply of getting a more accurate IQ by testing under more optimal conditions. Part of creating more optimal conditions in the case of disadvantaged children consists of giving at least two tests, the first only for practice and for letting the child get to know the examiner. I would put very little confidence in a single test score, especially if it is the child's first test and more especially if the child is from a poor background and of a different race from the examiner. But I also believe it is possible to obtain accurate assessments of a child's ability, and I would urge that attempts to evaluate preschool enrichment programs measure the gains against initially valid scores. If there is not evidence that this precaution has been taken, and if there is no control group, one might as well subtract at least 5 points from the gain scores as having little or nothing to do with real intellectual growth.

It is interesting that the IQ gains typically found in enrichment programs are of about the same magnitude and durability as those found in studies of the effects of direct coaching and practice on intelligence tests. The average IQ gain in such studies is about nine or ten points (Vernon, 1954).

What Is Really Changed When We Boost IQ? Test scores may increase after special educational treatment, but one must then ask which components of test variance account for the gain. Is it *g* that gains, or is it something less central to our concept of intelligence? We will not know for sure until someone does a factor analysis of pre- and post-test scores, including a number of "reference" tests that were not a part of the pre-test battery. We should also factor analyze the tests at the item level, to see which types of test items reflect the most gain. Are they the items with the highest cultural loadings? It is worth noting that the studies showing authentic gains used tests which are relatively high in cultural loading. I have found no studies that demonstrated gains in relatively noncultural or nonverbal tests like Cattell's Culture Fair Tests and Raven's Progressive Matrices.

Furthermore, if gain consists of actual improvement in cognitive skills rather than of acquisition of simple information, it must be asked whether the gain in skill represents the intellectual skill that the test normally measures, and which, because of the test's high heritability, presumably reflects some important, biologically based aspect of mental development. Let me cite one example. In a well-known experiment Gates and Taylor (1925) gave young children daily practice over several months in repeating auditory digit series, just like the digit span subtests in the Wechsler and Stanford-Binet. The practice resulted in a marked gain in the children's digit span, equivalent to an IQ gain of about 20 points. But when the children were retested after an interval of six months without practicing digit recall, their digit performance was precisely at the level expected for their mental age as determined by other tests. The gains had been lost, and the digit test once again accurately reflected the children's overall level of mental development, as it did before the practice period. The well-known later "fading" of IQ gains acquired early in enrichment programs may be a similar phenomenon.

But there is another phenomenon that probably is even more important as one of the factors working against the persistence of initial gains. This is the so-called "cumulative deficit" phenomenon, the fact that many children called disadvantaged show a decline in IQ from preschool age through at least elementary school. The term "cumulative deficit" may not be inappropriate in its connotations with respect to scholastic attainment, but it is probably a misleading misnomer when applied to the normal negatively accelerated growth rate of developmental characteristics of intelligence. The same phenomenon can be seen in growth curves of stature, but no one would refer to the fact that some children gain height at a slower rate and level off at a lower asymptote as a "cumulative deficit." In short, it seems likely that some of the loss in initial gains is due to the more negatively accelerated growth curve for intelligence in disadvantaged children and is not necessarily due to waning or discontinuance of the instructional effort. The effort required to boost IQ from 80 to 90 at 4 or 5 years of age is miniscule compared to the effort that would be required by age 9 or 10. "Gains" for experimental children in this range, in fact, take the form of superiority over a control group which has declined in IQ; the "enriched" group is simply prevented from falling behind, so there is no absolute gain in IQ, but only an advantage relative to a declining control group. Because of the apparently ephemeral nature of the gains seen in preschool programs, judgments of these programs' effectiveness in making a significant impact on intellectual development should be based on long range results.

A further step in proving the effectiveness of a particular program is to demonstrate that it can be applied with comparable success by other individuals in other schools, and, if it is to be practicable on a large scale, to determine if it works in the hands of somewhat less inspired and less dedicated practitioners than the few who originated it or first put it into practice on a small scale. As an example of what can happen when a small-scale project gets translated to a large-scale one, we can note Kenneth B. Clark's (1963, p. 160) enthusiastic and optimistic description of a "total push" intensive compensatory program with originated in one school serving disadvantaged children in New York City, with initially encouraging results. Clark said, "these positive results can be duplicated in every school of this type." In fact, it was tried in 40 other New York schools, and became known as the Higher Horizons program. After three years of the program the children in it showed no gains whatever and even averaged slightly lower in achievement and IQ than similar children in ordinary schools (U.S. Commission on Civil Rights, 1967, p. 125).

Finally, little is known about the range of IQ most likely to show genuine gains under enrichment. None of the data I have seen in this area permits any clear judgment on this matter. It would be unwarranted to assume at this time that special educational programs push the whole IQ distribution up the scale, so that, for example, they would yield a higher percentage of children with IQs higher than two standard deviations above the mean. After a "total push" program, IQs, if they change at all, may no longer be normally distributed, so that the gains would not much affect the frequencies at the tails of the distribution. We simply do not know the answer to this at present, since the relevant data are lacking.

Hothouse or Fertilizer? There seems to be little doubt that a deprived environment can stunt intellectual development and that immersion in a good environment in early childhood can largely overcome the effects of deprivation, permitting the individual's genetic potential to be reflected in his performance. But can special enrichment and instructional procedures go beyond the prevention or amelioration of stunting? As Vandenberg (1968, p. 49) has asked, does enrichment act in a manner similar to a hothouse, forcing an early bloom which is nevertheless no different from a normal bloom, or does it act more like a *fertilizer,* producing bigger and better yields? There can be little question about the hothouse aspect of early stimulation and instruction. Within limits, children can learn many things at an earlier age than that at which they are normally taught in school. This is especially true of forms of associative learning which are mainly a function of time spent in the learning activity rather than of the development of more complex cognitive structures. While most children, for example, do not learn the alphabet until 5 or 6 years of age, they are fully capable of doing so at about 3, but it simply requires more time spent in learning. The cognitive structures involved are relatively simple as compared with, say, learning to copy a triangle or a diamond. Teaching a 3-year-old to copy a diamond is practically impossible; at five it is extremely difficult; at seven the child apparently needs no "teaching"—he copies the diamond easily. And the child of five who has been *taught* to copy the diamond seems to have learned something different from what the seven-year-old "knows" who can do it without being "taught." Though the final performance of the five-year-old and the seven-year-old may *look* alike, we know that the cognitive structures underlying their performance are different. Certain basic skills can be acquired either associatively by rote learning or cognitively by conceptual learning, and what superficially may appear to be the same performance may be acquired in preschoolers at an associative level, while at a conceptual level in older children. Both the four-year-old and the six-year-old may know that 2+2=4, but this knowledge can be associative or cognitive. Insufficient attention has been given in preschool programs so far to the shift from associative to cognitive learning. The preschooler's capacity for associative learning is already quite well developed, but his cognitive or conceptual capacities are as yet rudimentary and will undergo their period of most rapid change between about five and seven years of age (White, 1965). We need to know more about what children can learn before age five that will transfer positively to later learning. Does learning something on an associative level facilitate or hinder learning the same content on a conceptual level?

While some preschool and compensatory programs have demonstrated earlier than normal learning of certain skills, the evidence for accelerative cognitive development or the speed of learning is practically nil. But usually this distinction is not made between sheer performance and the nature of the cognitive structures which support the gains in performance, and so the research leaves the issue in doubt. The answer to such questions is to be found in the study of the kinds and amount of transfer that result from some specific learning. The capacity for transfer of training is one of the essential aspects of what we mean by intelligence. The IQ gains reported in enrichment studies appear to be gains more in what Cattell calls "crystallized," in contract to "fluid," intelligence. This is not to say that gains of this type are not highly worthwhile. But having a clearer conception of just what the gains consist of will give us a better idea of how they can be most effectively followed up and of what can be expected of their effects on later learning and achievement.

Specific Programs. Hodges and Spicker (1967) have summarized a number of the more substantial preschool intervention studies designed to improve the intellectual capabilities and scholastic success of disadvantaged children. Here are some typical examples.

The *Indiana Project* focused on deprived Appalachian white children five years of age, with IQs in the range of 50 to 85. The children spent one year in a special kindergarten with a structured program designed to remedy specific diagnosed deficiencies of individual children in the areas of language development, fine motor coordination, concept formation, and socialization. Evaluation extended over two years, and gains were measured against three control groups: regular kindergarten, children who stayed at home during the kindergarten year, and children at home in another similar community. The average gain (measured against all three controls) after two years was 10.8 IQ points on the Stanford-Binet (final IQ 97.4) and 4.0 IQ points on the Peabody Picture Vocabulary Test (final IQ 90.4).

The *Perry Preschool Project* at Ypsilanti, Michigan, also was directed at disadvantaged preschool children with IQs between 50 and 85. The program was aimed at remedying lacks largely in the verbal prerequisites for first-grade learning and involved the parents as well as the children. There was a significant gain of 8.9 IQ points in the Stanford-Binet after one year of the preschool, but by the end of the second grade, the experimental group exceeded the controls, who had had no preschool attendance, by only 1.6 IQ points, a nonsignificant gain.

The *Early Training Project* under the direction of Gray and Klaus at Peabody College is described as a multiple intervention program, meaning that it included not only preschool enrichment but work with the disadvantaged children's mothers to increase their ability to stimulate their child's cognitive development at home. Two experimental groups, with two and three summers of preschool enrichment experience in a special school plus home visits by the training staff, experienced an average gain, four years after the start of the program, of 7.2 IQ points over a control group on the Stanford-Binet (final IQ of E group was 93.6).

The *Durham Education Improvement Program* (1966-1967b) has focused on preschool children from impoverished homes. The basic assumption of the program is stated as follows: "First, Durham's disadvantaged youngsters are considered normal at birth and potentially normal academic achievers, though they are frequently subjected to conditions jeopardizing their physical and emotional health. It is further assumed that they adapt to their environment according to the same laws of learning which apply to all children." The program is one of the most comprehensive and intensive efforts yet made to improve the educability of children from backgrounds of poverty. The IQ gains over about an eight to nine months' interval for various groups of preschoolers in the program are raw pre-post test gains, not gains over a control group. The average IQ gains on three different tests were 5.32 (Peabody Picture Vocabulary), 2.62 (Stanford-Binet), and 9.27 (Wechsler Intelligence Scale for Children). In most cases, IQ's changed from the 80s to the 90s.

The well-known Bereiter-Engelmann (1966) program at the University of Illinois is probably the most sharply focused of all. It aims not at all-round enrichment of the child's experience but at teaching specific cognitive skills, particularly of a logical, semantic nature (as contrasted with more diffuse "verbal stimulation"). The emphasis is on information processing skills considered essential for school learning.

The Bereiter-Engelmann preschool is said to be academically oriented, since each day throughout the school year the children receive twenty-minute periods of intensive instruction in three major content areas—language, reading, and arithmetic. The instruction, in small groups, explicitly involves maintaining a high level of attention, motivation, and participation from every child. Overt and emphatic repetition by the children are important ingredients in the instructional process. The pre-post gains (not measured against a control group) in Stanford-Binet IQ over an eighteen months' period are about 8 to 10 points. Larger gains are shown in tests that have clearly identifiable content which can reflect the areas receiving specific instruction, such as the Illinois Test of Psycholinguistic Abilities and tests of reading and arithmetic (Bereiter & Engelmann, 1968). The authors note that the gains are shared about equally by all children.

Bereiter and Engelmann, correctly, I believe, put less stock in the IQ gains than in the gains in scholastic performance achieved by the children in their program. They comment that the children's IQs were still remarkably low for children who performed at the academic level actually attained in the program. Their scholastic performance was commensurate with that of children 10 or 20 points higher in IQ. Such is the advantage of highly focused training—it can significantly boost the basic skills that count most. Bereiter and Engelmann (1966, p. 54) comment, "...to have taught children in a two-hour period per day enough over a broad area to bring the average IQ up to 110 of 120 would have been an impossibility." An important point of the Bereiter-Engelmann program is that it shows that scholastic performance—the acquisition of the basic skills—can be boosted much more, at least in the early years, than can the IQ, and that highly concentrated, direct instruction is more effective than the more diffuse cultural enrichment.

The largest IQ gains I have seen and for which I was also able to examine the data and statistical analyses were reported by Karnes (1968), whose preschool program at the University of Illinois is based on an intensive attempt to ameliorate specific learning deficits in disadvantaged three-year-old children. Between the average age of 3 years 3 months and 4 years 1 month, children in the program showed a gain of 16.9 points in the Stanford-Binet, which a control group showed a loss of 2.8 over the same period, making for a net gain of 19.7 IQ points for the experimental group. Despite rather small samples (E=15, C=14), this gain is highly significant statistically (a probability of less than 1 in 1000 of occurring by chance). Even so, I believe such findings need to be replicated for proper evaluation, and the durability of the gains needs to be assessed by follow-up studies over the years. There remains the question of the extent to which specific learning at age three affects cognitive structures which normally do not emerge until six or seven years of age and whether induced gains at an early level of mental development show appreciable "transfer" to later stages. It is hoped that investigators can keep sufficient track of children in preschool programs to permit a later follow-up which could answer these questions. An initial small sample size mitigates against this possibility, and so proper research programs should be planned accordingly.

"Expectancy Gain." Do disadvantaged children perform relatively poorly on intelligence tests because their teachers have low expectations for their ability? This belief has gained popular currency through an experiment by Rosenthal and Jacobson (1968). Their notion is that the teacher's expectations for the child's performance act as a self-fulfilling prophecy. Consequently, according to this hypothesis, one way to boost these children's intelligence, and presumably their general scholastic performance as well, is to cause teachers to hold out higher expectations of these children's ability. To test this idea, Rosenthal and Jacobson picked about five children at random from each of the classes in an elementary school and then informed the classroom teachers that, according to test results, the selected children were expected to show unusually intellectual gains in the coming year. Since the "high expectancy" children in each class were actually selected at random, the only way they differed from their classmates was presumably in the minds of their teachers. Group IQ tests administered by the teachers on three occasions during the school year showed a significantly larger gain in the "high expectancy" children than in their classmates. Both groups gained in IQ by amounts that are typically found as a result of direct coaching or of "total push" educational programs. Yet the authors note that "Nothing was done directly for the disadvantaged child at Oak School. There was no crash program to improve his reading ability, no special lesson plans, no extra time for tutoring, no trips to museums or art galleries. There was only the belief that the children bore watching, that they had intellectual competencies that would in due course be revealed" (p. 181). The net total IQ gain (i.e., Expectancy group minus Control group) for all grades was 3.8 points. Net gain in

verbal IQ was 2.1; for Reasoning (nonverbal) IQ the gain was 7.2. Differences were largest in grades 1 and 2 and became negligible in higher grades. The statistical significance of the gains is open to question and permits no clear-cut conclusions. (The estimation of the error variance is at issue: the investigators emphasized the individual pupil's scores as the unit of analysis rather than the means of the E and C groups for each classroom as the unit. The latter procedure, which is regarded as more rigorous by many statisticians, yields statistically negligible results.)

Because of the questionable statistical significance of the results of this study, there may actually be no phenomenon that needs to be explained. Other questionable aspects of the conduct of the experiment make it mandatory that the results be replicated under better conditions before any conclusions from the study be taken seriously or used as a basis for educational policy. For example, the same form of the group-administered IQ test was used for each testing, so that specific practice gains were maximized. The teachers themselves administered the tests, which is a faux pas par excellence in research of this type. The dependability of teacher-administered group tests leaves much to be desired. Would any gains beyond those normally expected from general test familiarity have been found if the children's IQs had been accurately measured in the first place by individual tests administered by qualified psychometrists without knowledge of the purpose of the experiment? These are some of the conditions under which such an experiment must be conducted if it is to inspire any confidence in its results.

Conclusions About IQ Gains. The evidence so far suggests the tentative conclusion that the pay-off of preschool and compensatory programs in terms of IQ gains is small. Greater gains are possible in scholastic performance when instructional techniques are intensive and highly focused, as in the Bereiter-Engelmann program. Educators would probably do better to concern themselves with teaching basic skills directly than with attempting to boost overall cognitive development. By the same token, they should de-emphasize IQ tests as a means of assessing gains, and use mainly direct tests of the skills the instructional program is intended to inculcate. The techniques for raising intelligence per se, in the sense of *g*, probably lie more in the province of the biological sciences than in psychology and education.

Gorden and Wilkerson (1966, pp. 158-159) have made what seems to me perhaps the wisest statement I have encountered regarding the proper aims of intervention programs:

> ...the unexpressed purpose of most compensatory programs is to make disadvantaged children as much as possible like the kinds of children with whom the school has been successful, and our standard of educational success is how well they approximate middle-class children in school performance. It is not at all clear that the concept of compensatory education is the one which will most appropriately meet the problems of the disadvantaged. These children are *not* middle-class children, many of them never *will* be, and they can never be anything but second-rate as long as they are thought of as potentially middle-class children.... At best they are different, and an approach which views this difference merely as something to be overcome is probably doomed to failure.

"Learning Quotient" versus Intelligence Quotient

If many of the children called culturally disadvantaged are indeed "different" in ways that have educational implications, we must learn as much as possible about the real nature of these differences. To what extent do the differences consist of more than just the well known differences in IQ and scholastic achievement, and, of course, the obvious differences in cultural advantages in the home?

Evidence is now emerging that there are stable ethnic differences in *patterns* of ability and that these patterns are invariant across wide socioeconomic differences (Lesser, Fifer, & Clark, 1965; Stodolsky & Lesser, 1967). Middle-class and lower-class groups differed about one standard deviation on all four abilities (Verbal, Reasoning, Number, Space) measured by Lesser and his co-workers, but the profile or pattern of scores was distinctively different for Chinese, Jewish, Negro, and Puerto Rican children, regardless of their social class. Such differences in patterns of ability are bound to interact with school instruction. The important question is how many other abilities there are that are not tapped by conventional tests for which there exist individual and group differences that interact with methods of instruction.

Through our research in Berkeley we are beginning to perceive what seems to be a very significant set of relationships with respect to patterns of ability which, unlike those of Lesser et al., seem to interact more with social class than with ethnic background.

In brief, we are finding that a unidimensional concept of intelligence is quite inadequate as a basis for understanding social class differences in ability. For example, the magnitude of test score differences between lower- and middle-class children does not always correspond to the apparent "cultural loading" of the test. Some of the least culturally loaded tests show the largest differences between lower and middle-class children. At least two dimensions must be postulated to comprehend the SES differences reported in the literature and found in our laboratory (see Jensen, 1968c, 1968d). These two dimensions and the hypothetical location of various test loadings on each dimension are shown in Figure 17. The horizontal axis represents the degree of cultural loading of the test. It is defined by the test's heritability. I have argued elsewhere (Jensen, 1968c) that the heritability index for a test is probably our best objective criterion of its culture-fairness. Just because tests do not stand at one or the other extreme of this continuum does not mean that the concept of culture-fairness is not useful in discussing psychological tests. The vertical axis in Figure 17 represents a continuum ranging from "simple" associative learning to complex cognitive or conceptual learning. I have hypothesized two genotypically distinct basic processes underlying this continuum, labeled Level I (associative ability) and Level II (conceptual ability). Level I involves the neural registration and consolidation of stimulus inputs and the formation of associations. There is relatively little transformation of the input, so there is a high correspondence between the forms of the stimulus input and the form of the response output. Level I ability is tapped mostly by tests such as digit memory, serial rote learning, selective trial-and-error learning with reinforcement (feedback) for correct responses, and in slightly less "pure" form by free recall of visually or verbally presented materials, and paired-association learning. Level II abilities, on the other hand, involve self-initiated elaboration and transformation of the stimulus input before it eventuates in an overt response. Concept learning and problem solving are good examples. The subject must actively manipulate the input to arrive at the output. This ability is best measured by intelligence tests with a low cultural loading and a high loading on *g*—for example, Raven's Progressive Matrices.

Social class differences in test performance are more strongly associated with the vertical dimension in Figure 17 than with the horizontal.

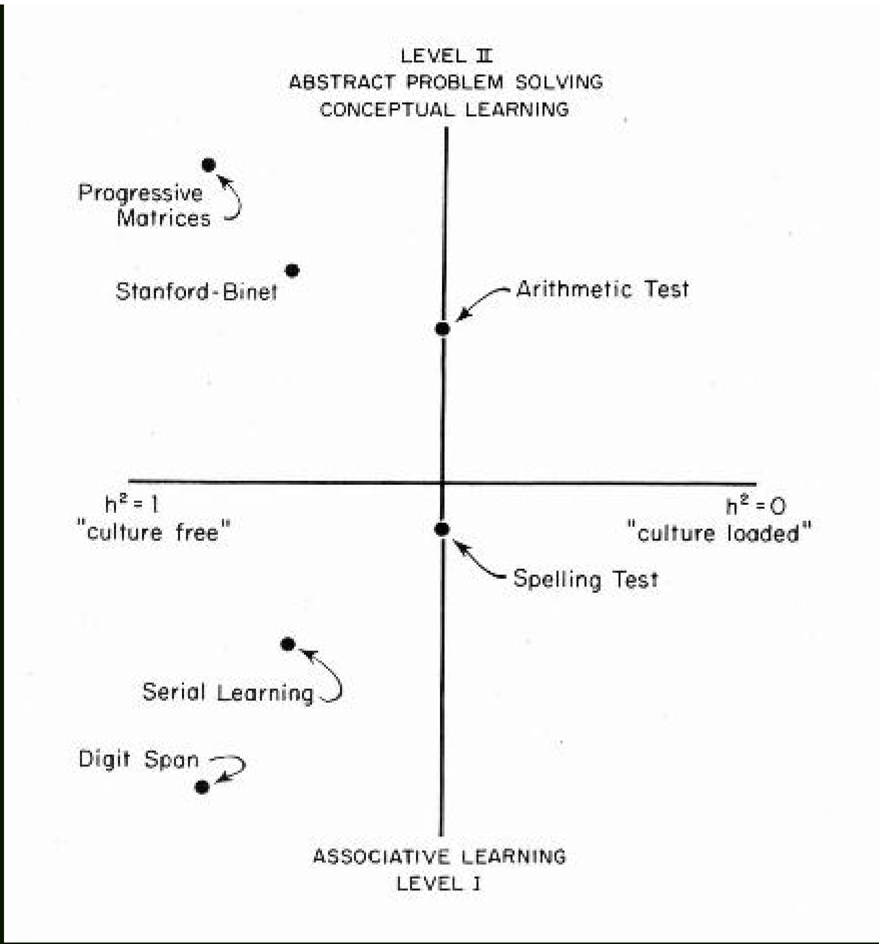

FIGURE 17.
The two-dimensional space required for comprehending social class differences in performance on tests of intelligence, learning ability, and scholastic achievement. The locations of the various "tests" are hypothetical.

Associative Learning Ability

Teachers of the disadvantaged have often remarked that many of these children seem much brighter than their IQs would lead one to expect, and that, even though their scholastic performance is usually as poor as that of middle-class children of similar IQ, the disadvantaged children usually appear much brighter in nonscholastic ways than do their middle-class counterparts in IQ. A lower-class child coming into a new class, for example, will learn the names of 20 or 30 children in a few days, will quickly pick up the rules and the know-how of various games on the playground, and so on—a kind of performance that would seem to belie his IQ, which may even be as low as 60. This gives the impression that the test is "unfair" to the disadvantaged child, since middle-class children in this range of IQ will spend a year in a classroom without learning the names of more than a few classmates, and they seem almost as inept on the playground and in social interaction as they are in their academic work.

We have objectified this observation by devising tests which can reveal these differences. The tests measure associative learning ability and show how fast a child can learn something relatively new and unfamiliar, right in the test situation. The child's performance does not depend primarily, as it would in conventional IQ tests, upon what he has already learned at home or elsewhere before he comes to take the test. We simply give him something to learn, under conditions which permit us to measure the rate and thoroughness of the learning. The tasks most frequently used are various forms of auditory digit memory, learning the serial order of a number of familiar objects or pictures of objects, learning to associate pairs of pictures of familiar objects, and free recall of names or objects presented from one to five times in a random order.

Our findings with these tests, which have been presented in greater detail elsewhere (Jensen, 1968a, 1968b, 1968d, 1968e; Jensen 1968f; Jensen & Rohwer, 1968), seem to me to be of great potential importance to the education of many of the children called disadvantaged. What we are finding, briefly, is this: lower-class children, whether white, Negro, or Mexican-American, perform as well on these direct learning tests as do middle-class children. Lower-class children in the IQ range of about 60 to 80 do markedly *better* than middle-class children who are in this range of IQ. Above about IQ 100, on the other hand, there is little or no difference between social class groups on the learning tests.

At first we thought we had finally discovered a measure of "culture-fair" testing, since we found no significant SES differences on these learning tests. But we can no longer reconcile this interpretation with all the facts now available. Some of the low SES children with low IQs on culturally loaded tests, like the Peabody Picture Vocabulary Tests, do very well on our learning tests, but do not have higher IQs on less culturally loaded tests of *g*, like the Progressive Matrices. It appears that we are dealing here with two kinds of abilities—associative learning ability (Level I) and cognitive or conceptual learning and problem-solving ability (Level II).

One particular test—free recall—shows the distinction quite well, since a slight variation in the test procedure makes the difference between whether it measures Level I or Level II. This is important, because it is sometimes claimed that low SES children do better on our learning tests than on IQ tests because the former are more interesting or more "relevant" to them, and thus make them more highly motivated to perform at their best. This is not a valid interpretation, since when essentially the same task is made either "associative" or "cognitive," we get differences of about one standard deviation in the mean scores of lower- and middle-class children. For example, 20 unrelated familiar objects (doll, toy car, comb, cup, etc.) are shown to children who are then asked to recall as many objects as they can in any order that may come to mind. The random presentation and recall are repeated five times to obtain a more reliable score. Lower- and middle-class elementary school children perform about the same on this task, although they differ some 15 to 20 points in IQ. This free recall test has a low correlation with IQ and the correlation is lower for the low SES children. But then we can change the recall test so that it gives quite different results.

This is shown in an experiment from our laboratory by Glasman (1968). (In this study SES and race are confounded, since the low SES group were Negro children and the middle SES group were white.) Again, 20 familiar objects are presented, but this time the objects are selected so that they can be classified into one of our categories, *animals, furniture, clothing,* or *foods.* There are five items in each of the four categories, but all 20 items are presented in a random order on each trial. Under this condition a larger social class difference shows up: the low SES children perform only slightly better on the average than they did on the uncategorized objects, while the middle SES children show a great improvement in performance which puts their scores about one standard deviation above the low SES children. Furthermore, there is much greater evidence of "clustering" the items in free recall for the middle SES than for the low SES children. That is, the middle-class children arrange the input in such a way that the order of the output in recall corresponds to the categories to which the objects may be assigned. The low SES children show less clustering in this fashion, although many show rather idiosyncratic pair-wise "clusters" that persist from trial to trial. There is a high correlation between the strength of the clustering tendency and the amount of recall. Also, clustering tendency is strongly related to age. Kindergartners, for example, show little difference between recall of categorized and uncategorized lists, and at this age SES differences in performance are nil. By fourth or fifth grade, however, the SES differences in clustering tendency are great, with a correspondingly large difference in ability to recall categorized lists.

It is interesting, also, that the recall of categorized lists correlated highly with IQ. In fact, when mental age or IQ is partialled out of the results, there are no significant remaining SES differences in recall. Post-test interviews showed that the recall differences for the two social class groups cannot be attributed to the low SES group's not knowing the category names. The children know the categories but tend not to use them spontaneously in recalling the list.

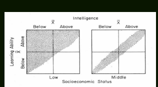

FIGURE 18.

Schematic illustration of the essential form of the correlation scatter-diagram for the relationship between associative learning ability and IQ in Low SES and Upper-Middle SES groups.

In general, we find that Level I associative learning tasks correlate very substantially with IQ among middle-class children but have very low correlations with IQ among lower-class children (Jensen, 1968b). The reason for this difference in correlations can be traced back to the form of the scatter diagrams for the middle and low SES groups, which is shown schematically in Figure 18. Since large representative samples of the entire school population have not been studied so far, the exact form of the correlation scatter diagram has not yet been well established, but the schematic portrayal of Figure 18 is what could be most reasonably hypothesized on the basis of several lines of evidence now available. (Data on a representative sample of 5000 children given Level I and Level II tests are now being analyzed to establish the forms of the correlation plots for low and middle SES groups.) The form of the correlation as it now appears suggests a hierarchical arrangement of mental abilities, such that Level I ability is necessary but not sufficient for Level II. That is, high performance on Level II tasks depends upon better than average ability on Level I, but the reverse does not hold. If this is true, the data can be understood in terms of one additional hypothesis, namely, that Level I ability is distributed about the same in all social class groups, while Level II ability is distributed differently in lower and middle SES groups. The hypothesis is expressed graphically in Figure 19. Heritability studies of Level II tests cause me to believe that Level II processes are not just the result of interaction between Level I learning ability and experimentally acquired strategies or learning sets. That learning is necessary for Level II no one doubts, but certain neural structures must also be available for Level II abilities to develop, and these are conceived of as being different from the neural structures underlying Level I. The genetic factors involved in each of these types of ability are presumed to have become differentially distributed in the population as a function of social class, since Level II has been most important for scholastic performance under the traditional methods of instruction.

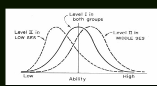

FIGURE 19.

Hypothetical distributions of Level I (solid line) and Level II (dashed line) abilities in middle-class and culturally disadvantaged populations.

From evidence on age differences in different tasks on the Level I-Level 2 continuum (e.g., Jensen & Rohwer, 1965), I have suggested one additional hypothesis concerning the developmental rates of Level I and Level II abilities in lower and middle SES group, as depicted in Figure 20. Level I abilities are seen as developing rapidly and as having about the same course of development and final level in both lower and middle SES groups. Level II abilities, by contrast, develop slowly at first, attain prominence between four and six years of age, and show an increasing difference between the SES groups with increasing age. This formulation is consistent with the increasing SES differences in mental age on standard IQ tests, which tap mostly Level II ability.

Thus, ordinary IQ tests are not seen as being "unfair" in the sense of yielding inaccurate or invalid measures for the many disadvantaged children who obtain low scores. If they are unfair, it is because they tap only one part of the total spectrum of mental abilities and do not reveal that aspect of mental ability which may be the disadvantaged child's strongest point—the ability for associative learning.

Since traditional methods of classroom instruction were evolved in populations having a predominantly middle-class pattern of abilities, they put great emphasis on cognitive learning rather than associative learning. And in the post-Sputnik era, education has seen an increased emphasis on cognitive and conceptual learning, much to the disadvantage of many children whose mode of learning is predominantly associative. Many of the basic skills can be learned by various means, and an educational system that puts inordinate emphasis on only one mode or style of learning will obtain meager results from the children who do not fit this pattern. At present, I believe that the educational system—even as it falteringly attempts to help the disadvantaged—operated in such a way as to maximize the importance of Level II (i.e., intelligence or *g*) as a source of variance in scholastic performance. Too often, if a child does not learn the school subject matter when taught in a way that depends largely on being average or above average on *g*, he does not learn at all, so that we find high school students who have failed to learn basic skills which they could easily have learned many years earlier by means that do not depend much on *g*. It may well be true that many children today are confronted in our schools with an educational philosophy and methodology which were mainly shaped in the past, entirely without any roots in these children's genetic and cultural heritage. The educational system was never allowed to evolve in such a way as to maximize the actual potential for learning that is latent in these children's patterns of abilities. If a child cannot show that he "understands" the meaning of 1+1=2 in some abstract, verbal, cognitive sense, he is, in effect, not allowed to go on to learn 2+2=4. I am reasonably convinced that all the basic scholastic skills can be learned by children with normal Level I learning ability, provided the instructional techniques do not make *g* (i.e., Level II) the *sine qua non* of being able to learn. Educational researchers must discover and devise teaching methods that capitalize on existing abilities for the acquisition of those basic skills which students will need in order to get good jobs when they leave school. I believe there will be greater rewards for all concerned if we further explore different types of abilities and modes of learning, and seek to discover how these various abilities can serve the aims of education. This seems more promising than acting as though only one pattern of abilities, emphasizing *g*, can succeed educationally, and therefore trying to inculcate this one ability pattern in all children.

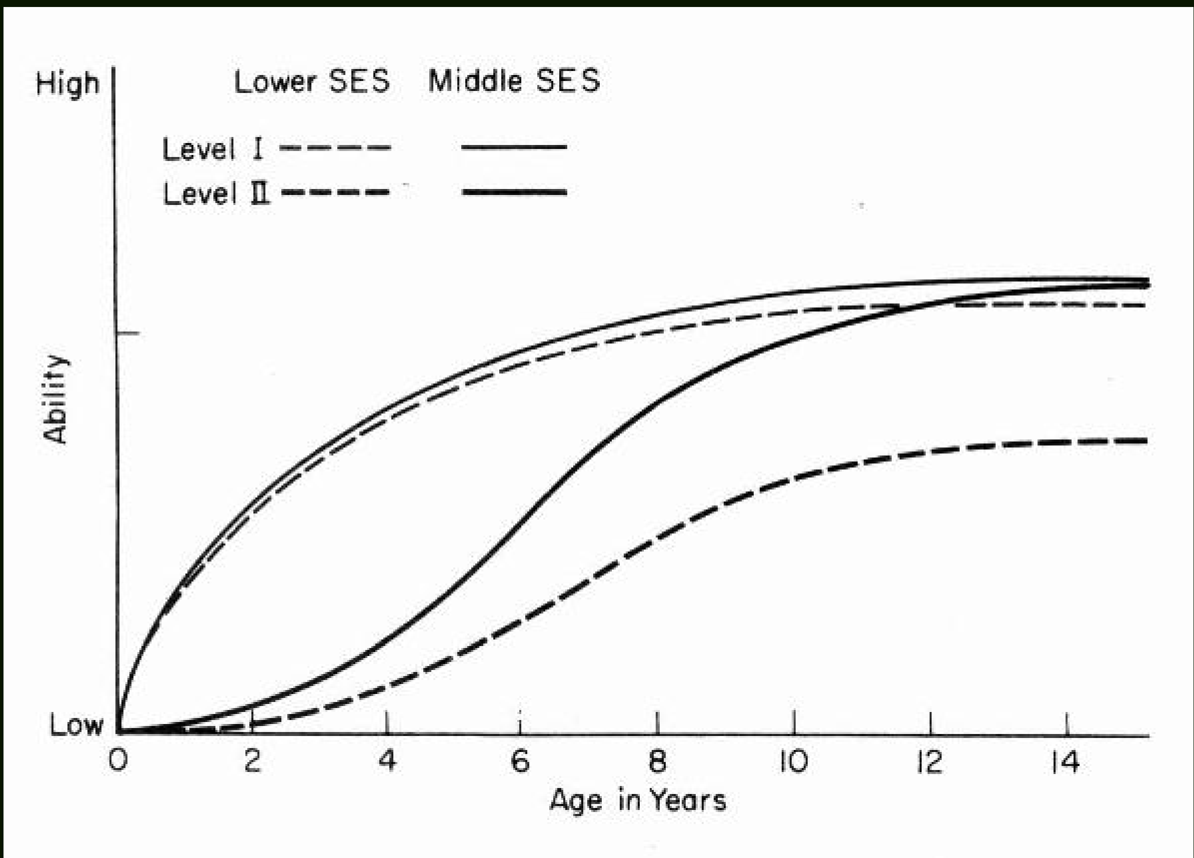

FIGURE 20.
Hypothetical growth curves for Level I and Level II abilities in middle SES and low SES populations.

If the theories I have briefly outlined here become fully substantiated, the next step will be to develop the techniques by which school learning can be most effectively achieved in accordance with different patterns of ability. By all means, schools must discover *g* wherever it exists and see to it that its educational correlates are fully encouraged and cultivated. There can be little doubt that certain educational and occupational attainments depend more upon *g* than upon another other single ability. But schools must also be able to find ways of utilizing other strengths in children whose major strength is not of the cognitive variety. One of the great and relatively untapped reservoirs of mental ability in the disadvantaged, it appears from our research, is the basic ability to learn. We can do more to marshal this strength for educational purposes.

If diversity of mental abilities, as of most other human characteristics, is a basic fact of nature, as the evidence indicates, and if the ideal of universal education is to be successfully pursued, it seems a reasonable conclusion that schools and society must provide a range and diversity of educational methods, programs, and goals, and of occupational opportunities, just as wide as the range of human abilities. Accordingly, the ideal of equality of educational opportunity should not be interpreted as uniformity of faculties, instructional techniques, and educational aims for all children. Diversity rather than uniformity of approaches and aims would seem to be the key to making education rewarding for children of different patterns of ability. The reality of individual differences thus need not mean education rewards for some children and frustration and defeat for others.

Arthur Jensen's Footnotes.

1. Some of the largest and most highly publicized programs of compensatory education that have been held up as models but which produced absolutely no significant improvement in the scholastic achievement of disadvantaged students are: the *Banneker Project* in St. Louis (8 years), *Higher Horizons* in New York (5 years), More Effective Schools in New York (3 years), and large-scale programs in Syracuse, Seattle, Philadelphia, Berkeley, and a score of other cities (for detailed reports see U.S. Commission on Civil Rights, 1967, pp. 115-140). Reports on Project Head Start indicate that initial gains of 5 to 10 points in IQ on conventional intelligence tests are a common finding, but this gain usually does not hold up through the first year of regular schooling. More positive claims for the efficacy of Head Start involve evidence of the detection and correction of mental disabilities in disadvantaged preschool children and the reportedly favorable effects of the program on children's self-confidence, motivation, and attitudes toward work.

2. I am indebted to Professor Otis Dudley Duncan (1968, pp. 80-100) for providing this information.

3. I am grateful to University of California geneticist Dr. Jack Lester King for making these calculations, which are based on the assumption that the heritability of IQ is .80, a value which is the average of all the major studies of the heritability of intelligence.

4. Heritability in the narrow sense is an estimate of the proportion of genic variance without consideration of dominance and epistasis. This contrasts with equation (3), the definition of H, which includes estimates for these two factors. Signified as H_N, heritability in the narrow sense is conceptually defined as:

$$H_N = (V_G + V_{AM}) / (V_p - V_e)$$

5. After this article went to press I received a personal communication from Professor Lloyd G. Humphreys who pointed out some arguments that indicate I may have underestimated the heritability of scholastic achievement and that its heritability may actually be considerably closer to the heritability of intelligence. The argument involves two main points: (1) the fact that some of the achievement tests that entered into the average estimate of heritability are tests of specific achievements, rather than omnibus achievement tests, and therefore would correspond more to the separate subscales of the usually intelligence tests, which are known to have somewhat lower heritabilities than the composite scores; and

(2) scores on some of the achievement tests are age-related, so that fraternal twin correlations, in relation to other kinship correlations, are unduly inflated by common factor of age. When age is partialled out of the MZ and DZ twin correlations, the estimate of heritability based on MZ and DZ twin comparisons is increased. However, an omnibus achievement test (Standard Achievement) yielding an overall Educational Age score had a heritability of only .46 (as compared with .63 for Stanford-Binet IQ and .70 for Otis IQ based on the same set of MZ and DZ twins), with age partialled out of the twin correlations (Newman, Freeman, and Holzinger, 1937, p. 97). Rank in high school graduating class, which is an overall index of scholastic performance and is little affected by age yields heritability coefficients below .40 in a nationwide sample (Nichols & Bilbro, 1966). The issue clearly needs further study, but the best conclusion that can be drawn from the existing evidence, I believe, still is that the heritability of scholastic achievement is less than for intelligence, but the amount of the difference cannot be precisely estimated at present.

6. The most comprehensive and sophisticated discussion of the genic-behavior analysis of race differences that I have found is by Spuhler and Lindzey (1967).

References

Altus, W.D. Birth order and its sequelae. *Science,* 19666, **151,** 44-59.

Anastasi, A. Intelligence and family size, *Psycho. Bull.,* 1956, **53,** 187-209.

Bejema, C.J. Estimation of the direction and intensity of natural selection in relation to human intelligence by means of the intrinsic rate of natural increase. *Eugen. Quart.,* 1963, **10,** 175-187.

Bejema, C.J. Relation of fertility to educational attainment in a Kalamazoo public school population: A follow-up study, *Eugen. Quart.,* 1966, **13,** 306-315.

Bayley, N. Research in child development: A longitudinal perspective. *Merrill-Palmer Quart. Behav. Developm.,* 1965, **11,** 183-208. (b)

Bayley, N. Comparisons of mental and motor test scores for ages 1-15 months by sex, birth order, race, geographical location, and education of parents. *Child Developm.,* 1965, **36,** 379-411. (a)

Bayley, N. Learning in adulthood: The role of intelligence. In H.J. Klausmeier & C.W. Harris (Eds.), *Analyses of concept learning.* New York: Academic Press, 1966. Pp. 117-138.

Bayley, N. Behavioral correlates of mental growth: Birth to thirty-six years. *Amer. Psychol.,* 1968, **23,** 1-17.

Bereiter, C., & Engelmann, S. *Teaching disadvantaged children in the preschool.* Englewood Cliffs, N.J.: Prentice-Hall, 1966.

Bereiter, C., & Engelmann, S. An academically oriented preschool for disadvantaged children: Results from the initial experimental group. In D.W. Brison & J. Hill (Eds.), *Psychology and early childhood education.* Ontario Institute for Studies in Education, 1968. No. 4. Pp. 17-36.

Bilodeau, E.A. (Ed.), *Acquisition of skill.* New York: Academic Press, 1966.

Brison, D.W. Can and should learning be accelerated? In D.W. Brison (Ed.), *Accelerated learning and fostering creativity.* Toronto, Canada: Ontario Institute for Studies in Education, 1968. Pp. 5-9.

Bronfenbrenner, U. The psychological costs of quality and equality in education. *Child Developm.,* 1967, **38,** 909-925.

Buck, C. Discussion of "Culturally related reproductive factors in mental retardation" by Graves et al. Paper read at Conference on Sociocultural Aspects of mental Retardation, Peabody College, Nashville, Tenn., June, 1968.

Burks, B.S. The relative influence of nature and nurture upon mental development: A comparative study of foster parent-foster child resemblance and true parent-true child resemblance. *Yearb. Nat. Soc. Stud. Educ.,* 1928, **27, (I),** 219-316.

Burt, C. The evidence for the concept of intelligence. *Brit. J. educ. Psychol.,* 1955, **25,** 158-177.

Burt, C. The distribution of intelligence. *Brit. J. Psychol.* 1957, **48,** 161-175.

Burt, C. The inheritance of mental ability. *Amer. Psychol.,* 1958, **13,** 1-15.

Burt, C. Class difference in general intelligence: III. *Brit. J. Stat. Psychol.,* 1959, **12,** 15-33.

Burt, C. Intelligence and social mobility. *Brit. J. Stat. Psychol.,* 1961, **14,** 3-24.

Burt, C. Is intelligence distributed normally? *Brit. J. Stat. Psychol.,* 1963, **16,** 175-190.

Burt, C. The genetic determination of differences in intelligence: A study of monozygotic twins reared together and apart. *Brit. J. Psychol.,* 1966, **57,** 135-153.

Burt, C. Mental capacity and its critics. *Bull. Brit. Psychol. Soc.,* 1968, **21,** 11-18.

Burt, C., & Howard, M. The multifactorial theory of inheritance and its application to intelligence. *Brit. J. Stat. Psychol.,* 1956, **9,** 95-131.

Burt, C., & Howard, M. The relative influence of heredity and environment on assessments of intelligence. *Brit. J. Stat. Psychol.,* 1957, **10,** 99-104.

Bloom, B.S. *Stability and change in human characteristics.* New York: Wiley, 1964.

Carter, C.O. Differential fertility by intelligence. In J.E. Meade & A.S. Parkes (Eds.), *Genetic and environmental factors in human ability.* New York: Plenum Press, 1966. Pp. 185-200.

Cattell, R.B. The multiple abstract variance analysis equations and solutions: For nature-nurture research on continuous variables. *Psychol. Rev.,* 1960, **67,** 353-372.

Cattell, R.B. Theory of fluid and crystallized intelligence: A critical experiment. *J. educ. Psychol.,* 1963, **54,** 1-22.

Churchill, J.A., Neff, J.W., & Caldwell, D.F. Birth weight and intelligence. *Obstetrics and gynecology,* 1966, **28,** 425-429.

Clark, K.B. Educational stimulation of racially disadvantaged children. In A.H. Passow (Ed.), *Education in depressed areas.* New York: Teachers College Press, Columbia Univer., 1963. Pp. 142-162.

Coleman, J.S., et al. *Equality of educational opportunity,* U.S. Dept. of Health, Education, & Welfare, 1966.

Cooper, R., & Zubek, J. Effects of enriched and restricted early environments on the learning ability of bright and dull rats. *Canad. J. Psychol.,* 1958, **12,** 159-164.

Cravioto, J. Malnutrition and behavioral development in the preschool child. *Pre-school child malnutrition.* National Health Science, Public., 1966, No. 1282.

Cravioto, J., De Licardie, E.R., & Birch, H.G. Nutrition, growth, and neurointegrative development: An experimental and ecologic study. *Pediatrics,* 1966, **38,** 319-372.

Davis, K. Final note on a case of extreme isolation. *Amer. J. Sociol.,* 1947, **57,** 432-457.

Deutsch, M., Katz, I., & Jensen, A.R. (Eds.) *Social class, race, and psychological development.* New York: Holt, Rinehart & Winston, 1968.

Dobzhansky, T. Genetic differences between people cannot be ignored. *Scientific Res.,* 1968, **3,** 32-33. (a)

Dobzhansky, T. On genetics, sociology, and politics. *Perspect. Biol. Med.,* 1968, **11** 544-554. (b)

Dreger, R.M., & Miller, K.S. Comparative psychological studies of Negroes and whites in the United States. *Psychol. Bull.,* 1960, **57,** 361-402.

Dreger, R.M., & Miller, K.S. Comparative psychological studies of Negroes and whites in the United States: 1959-1965. *Psychol. Bull.,* 1968 (*Monogr. Suppl.* 70, No. 3, Part 2).

Duncan, O.D. Is the intelligence of the general population declining? *Amer. sociol. Rev.,* 1952, **17,** 401-407.

Duncan, O.D., Featherman, D.L, & Duncan, B. Socioeconomic background and occupational achievement: Extensions of a basic model. Final Report, Project No. 5-0074 (EO-191) U.S. Dept. of Health, Education, and Welfare, Office of Education, Bureau of Research, May, 1968.

Durham Education Improvement Program, 1966-1967. (a)

Durham Education Improvement Program, Research, 1966-1967. (b)

Dustman, R.E., & Beck, E.C. The visually evoked potential in twins. *Electroenceph. clin. Neurophysiol.,* 1965, **19,** 570-575.

Eckland, B.K. Genetics and sociology: A reconsideration. *Amer. Soc. Rev.,* 1967, **32,** 173-194.

Eells, K., et al. *Intelligence and cultural differences.* Chicago: Univ. Chicago Press, 1951.

Erlenmeyer-Kimling, L., & Jarvik, L.F. Genetics and intelligence: a review. *Science,* 1963, **142,** 1477-1479.

Falconer, D.S. *An introduction to quantitative genetics.* New York: Ronald Press, 1960.

Freeman, R.A. Schools and the elusive 'average children' concept. *Wall Street Journal,* July 8, 1968, p. 12.

Fuller, J.L., & Thompson, W.R. *Behavior genetics.* New York: Wiley, 1960.

Gates, A.I., & Taylor, G.A. An experimental study of the nature of improvement resulting from practice in mental function. *J. educ. Psychol.,* 1925, **16,** 583-593.

Geber, M. The psycho-motor development of African children in the first year, and the influence of maternal behavior. *J. soc. Psychol.,* 1958, **47,** 185-195.

Geber, M., & Dean, R.F.A. The state of development of newborn African children. *Lancet,* 1957, 1216-

1219.

Ghiselli, E.E. The measurement of occupational aptitude. *Univer. of California Publ. in Psychol.,* Vol. 8, No. 2. Berkeley, Calif.: Univer. of California Press, 1955.

Glasman, L.D. A social-class comparison of conceptual processes in children's free recall. Unpublished doctoral dissertation. Univer. of California, 1968.

Goertzel, V., & Goertzel, M.G. *Cradles of eminence,* London: Constable, 1962.

Goodenough, F.L. New evidence on environmental influence on intelligence. *Yearb. Nat. Soc. Stud. Educ.,* 1940, **39,** Part I, 307-365.

Gorden, E.W., & Wilkerson, D.A. *Compensatory education for the disadvantaged.* New York: College Entrance Examination Board, 1966.

Gottesman, I. Genetic aspects of intelligent behavior. In N.R. Ellis (Ed.), *Handbook of mental deficiency.* New York: McGraw-Hill, 1963. Pp. 253-296.

Gottesman, I. Biogenetics of race and class. In M. Deutsch, I. Katz, & A.R. Jensen (Eds.), *Social class, race, and psychological development.* New York: Holt, Rinehart & Winston, 1968. Pp. 11-51.

Graves, W.L., Freeman, M.G., & Thompson, J.D. Culturally related reproductive factors in mental retardation. Paper read at Conference on Sociocultural Aspects of Mental Retardation, Peabody College, Nashville, Tenn., June, 1968.

Guilford, J.P. *The nature of human intelligence.* New York: McGraw-Hill, 1967.

Hardy, J.B. Perinatal factors and intelligence. In S.F. Osler & R.E. Cooke (Eds.), *The biosocial basis of mental retardation.* Baltimore, Md.: The Johns Hopkins Press, 1965. Pp. 35-60.

Harlow, H.F., & Griffin, G. Induced mental and social deficits in Rhesus monkeys. In S.F. Osler & R.E. Cooke (Eds.), *The biosocial basis of mental retardation.* Baltimore, Md.: The Johns Hopkins Press, 1965. Pp. 87-106

Harlow, H.F., & Harlow, M.K. The mind of man. In *Yearbook of science and technology.* New York: McGraw-Hill, 1962.

Harrell, R.F., Woodyard, E. & Gates A.I. *The effect of mothers' diets on the intelligence of offspring.* New York: Bureau of Publications, Teachers College, 1955.

Heber, R. Research on education and habilitation of the mentally retarded. Paper read at Conference on Sociocultural Aspects of Mental Retardation, Peabody College, Nashville, Tenn., June, 1968.

Heber, R., Dever, R. & Conry, J. The influence of environmental and genetic variables on intellectual development. In H.J. Prehm, L.A. Hamerlynk, & J.E. Crosson (Eds.), *Behavioral Research in mental retardation.* Eugene, Oregon: Univer. Of Oregon Press, 1968. Pp. 1-25.

Heyns, O.S. *Abdominal decompression.* Johannesburg: Witwatersrand Univer. Press, 1963.

Higgins, J., Reed, S., & Reed, E. Intelligence and family size: A paradox resolved. *Eugen. Quart.,* 1962, **9,** 84-90.

Hill, A.C., & Jaffee, F.S. Negro fertility and family size preferences. In T. Parsons & K.B. Clark (Eds.), *The Negro American.* Cambridge, Mass.: Houghton-Mifflin, 1966. Pp. 134-159.

Hodges, W.L., & Spicker, H.H. The effects of preschool experiences on culturally deprived children. In W.W. Hartup & N.L Smothergill (Eds.), *The young child: Reviews of research.* Washington, D.C.: National Association for the Education of Young Children, 1967. Pp. 262-289.

Honzik, M.P. Developmental studies of parent-child resemblance in intelligence. *Child Developm.,* 1957, **28,** 215-228.

Honzik, M.P. the mental and motor test performances of infants diagnosed or suspected of brain injury. Final Rep., Contact SA 43 PH 2426. Washington, D.C.: National Institute of Health, National Institute of Neurological Disease and Blindness, Collaborative Research, May, 1962.

Hunt, J. McV. *Intelligence and Experience.* New York: Ronald Press, 1961.

Huntley, R.M.C. Heritability of intelligence. In J.E. Meade & A.S. Parker (Eds.), *Genetic and environmental factors in human ability.* New York: Plenum Press, 1966. Pp. 201-218.

Husén, T. Abilities of twins. *Scand. J. Psychol.,* 1960, **1,** 125-135.

Jensen, A.R. Estimation of the limits of heritability of traits by comparison of monozygotic and dizogotic twins. *Proc. Nat. Acad. Sci.,* 1967, **58,** 149-157. (a)

Jensen, A.R. Varieties of individual differences in learning. In Gagné, R.M. (Ed.), *Learning and individual differences.* Columbus, Ohio: Merrill, 1967. (b)

Jensen, A.R. Social class, race, and genetics: Implications for education. *Amer. Educ. Res. J.,* 1968, **5,** 1-42. (a)

Jensen, A.R. Patterns of mental ability and socioeconomic status. *Proc. Nat. Acad. Sci.,* 1968, **60,** 1330-1337. (b)

Jensen, A.R. Another look at culture-fair tests. *In Western Regional Conference on Testing Problems, Proceedings for 1968,* "Measurement for Educational Planning." Berkeley, Calif.: Educational Testing Service, Western Office, 1968. Pp. 50-104. (c)

Jensen, A.R. Intelligence, learning ability, and socioeconomic status. *J. spec. Educ.,* 1968. (d)

Jensen, A.R. Social class and verbal learning. In M. Deutsch, I. Katz, & A.R. Jensen (Eds.), *Social class, race, and psychological development.* New York: Holt, Rinehart & Winston, 1968. Pp. 115-174. (e)

Jensen, A.R. The culturally disadvantaged and the heredity-environment uncertainty. In J. Helmuth (Ed.), *The culturally disadvantaged child.* Vol. 2. Seattle, Wash.: Special Child Publications, 1968. (f)

Jensen, A.R. & Rohwer, W.D., Jr. Syntactical mediation of serial and paired-associate learning as a function of age. *Child Developm.,* 1965, **36,** 601-608.

Jensen, A.R., & Rohwer, W.D.,Jr. Mental retardation, mental age, and learning rate. *J. educ. Psychol.,* 1968.

Jepsen, N.P., & Bredmose, G.V. Investigations into the age of mentally deficient women at their first delivery. *Acta Psychiat. Scand.,* 1956 (Monogr. Suppl. 108), Pp. 203-210.

Jones, H.E. The environment and mental development. In L. Carmichael (Ed.), *Manual of child psychology.* (2nd ed.) New York: Wiley, 1954. Pp. 631-696.

Karnes, M.B. A research program to determine the effects of various preschool intervention programs on the development of disadvantaged children and the strategic age for such intervention. Paper read at

Amer. Educ. Res. Ass., Chicago, Feb., 1968.

Kempthorne, O. *An introduction to genetic statistics.* New York: Wiley, 1957.

Kennedy, W.A., Van De Riet, V., & White, J.C., Jr. A normative sample of intelligence and achievement of Negro elementary school children in the Southeastern United States. *Monogr. Soc. Res. Child Developm.,* 1963, **28,** No. 6.

Kushlick, A. Assessing the size of the problem of subnormality. In J.E. Meade & A.S. Parkes (Eds.), *Genetic and environmental factors in human ability.* New York: Plenum Press, 1966, Pp. 121-147.

Kuttner, R.E. *Biochemical anthropology.* In R.E. Kuttner (Ed.), *Race and modern science.* New York: Social Science Press, 1967. Pp. 197-222.

Kuttner, R.E. Letters to and from the editor. *Perpect. Biol. Med.,* 1968, **11,** 707-709.

Lawrence, E.M. An investigation into the relation between intelligence and inheritance. *Brit. J. Psychol. Monogr. Suppl.,* 1931, **16,** No. 5.

Leahy, A.M. Nature-nurture and intelligence. *Genet. Psychol. Monogr.,* 1935, **17,** 241-305.

Lesser, G.S., Fifer, G., & Clark, D.H. Mental abiltiies of children from different social-class and cultural groups, *Monogr. Soc. Res. Child Developm.,* 1965, **30,** (4)

Lindzey, G. Some remarks concerning incest, the incest taboo, and psychoanalytic theory. *American Psychol.,* 1967, **22,** 1051-1059.

Loehlin, J.C. Psychological genetics, from the study of human behavior. In R.B. Cattell (Ed.), *Handbook of modern personality theory.* New York: Aldine, in press.

Medical World News. Using speed of brain waves to test IQ. 1968, **9,** 26.

Mitra, S. Income, socioeconomic status, and fertility in the United States. *Eugen. Quart.,* 1966, **13,** 223-230.

Money, J. Two cytogenetic syndromes: Psychological comparisons 1. Intelligence and specific-factor quotients. *J. psychiat. Res.,* 1964, **2,** 223-231.

Moynihan, D.P. *The Negro family.* Washington, D.C.: Office of Policy Planning and Research, United States Department of Labor, 1965.

Moynihan, D.P. Employment, income, and the ordeal of the Negro family. In T. Parsons & K.B. Clark (Eds.), *The Negro American,* Cambridge, Mass.: Houghton-Mifflin, 1966. Pp. 134-159.

National Academy of Sciences. Racial studies: Academy states position on call for new research. *Science,* 1967, **158,** 892-893.

Naylor, A.E. & Myrianthopoulos, N.C. The relation of ethnic and selected socioeconomic factors to human birth-weight. *Ann. Hum. Genet.,* 1967, **31,** 71-83.

Nelson, G.K., & Dean, R.F.A. *Bull. World Health Organ.,* 1959, **21,** 779. Cited by G. Cravioto, Malnutrition and behavioral development in the preschool child. *Pre-school child malnutrition.* National Health Science, 1966, Public. No. 1282.

Newman, H.H., Freeman, F.N., & Holzinger, K.J. *Twins: A study of heredity and environment.* Chicago: Univ. of Chicago Press, 1937.

Nichols, R.C., & Bilbro, W.C., Jr. The diagnosis of twin zygosity. *Acta genet.,* 1966, **16,** 265-275.

Pettigrew, T. *A profile of the Negro American.* Princeton, N.J.: Van Nostrand, 1964.

Powledge, F. *To change a child—A report on the Institute for Developmental Studies.* Chicago: Quadrangle Books, 1967.

Reed, E.W., & Reed, S.C. *Mental retardation: A family study.* Philadelphia: W.B. Saunders Co., 1965.

Research Profile No. 11. Summary of progress in childhood disorders of the brain and nervous system. Washington, D.C.: Public Health Service, 1965.

Reymert, M.L., & Hinton, R.T., Jr. The effect of a change to a relatively superior environment upon the IQs of one hundred children. *yearb. Nat. Soc. Stud. Educ.,* 1940, **39, (I),** 255-268.

Rimland, B. *Infantile autism.* New York: Appleton-Century-Crofts, 1964.

Roberts, J.A.F. The genetics of mental deficiency. *Eugen. Rev.,* 1952, **44,** 71-83.

Roberts, R.C. some concepts and methods in quantitative genetics. In J. Hirsch (Ed.), *Behavior-genetic analysis.* New York: McGraw-Hill, 1967. Pp. 214-257.

Rosenthal, R., & Jacobson, L. *Pygmalion in the classroom.* New York: Holt, Rinehart & Winston, 1968.

Schull, W.J., & Neel, J.V. *The effects of inbreeding on Japanese children.* New York: Harper & Row, 1965.

Schwebel, M. *Who can be educated?* New York: Grove, 1968.

Scott, J.P., & Fuller, J.L. *Genetics and the social behavior of the dog.* Chicago: Univer. of Chicago Press, 1965.

Stodolsky, S.S., & Lesser, G. Learning patterns in the disadvantaged. *Harvard educ. Rev.,* 1967, **37,** 546-593.

Scrimshaw, N.S. Infant malnutrition and adult learning. *Saturday Review,* March 16, 1968. p. 64.

Shields, J. *Monozygotic twins brought up apart and brought up together.* London: Oxford Univer. Press, 1962.

Shuey, A.M. *The testing of Negro intelligence.* (2nd ed.) New York: Social Science Press, 1966.

Skeels, H.M. Adult status of children with contrasting early life experiences: A follow-up study. *Child Developm. Monogr.,* 1966, **31,** No. 3, Serial No. 105.

Skeels, H.M., & Dye, H.B. A study of the effects of differential stimulation on mentally retarded children. *Proc. Addr. Amer. Ass. Ment. Defic.,* 1939, **44,** 114-136.

Spuhler, J.N., & Lindzey, G. Racial differences in behavior. In J. Hirsch (Ed.), *Behavior-genetic analysis.* New York: McGraw-Hill, 1967. Pp. 366-414.

Stoch, M.B., & Smythe, P.M. Does undernutrition during infancy inhibit brain growth and subsequent intellectual development? *Arch. Dis. Childh.,* 1963, **38,** 546-552.

Stoddard, G.D. *The meaning of intelligence.* New York: MacMillan, 1943.

Stott, D.H. Interaction of heredity and environment in regard to 'measured intelligence.' *Brit. J. educ. Psychol.,* 1960, **30,** 95-102.

Stott, D.H. *Studies of troublesome children.* New York: Humanities Press, 1966.

Thompson, W.R. The inheritance and development of intelligence. *Res. Pub. Ass. Nerv. Ment. Dis.,* 1954, **33,** 209-331.

Thorndike, E.L. Measurement of twins. *J. Philos., Psychol., Sci. Meth.,* 1905, **2,** 547-553.

Tuddenham, R.D. Psychometricizing Piaget's méthode clinique. Paper read at Amer. Educ. Res. Ass., Chicago, February, 1968.

Tyler, L.E. *The psychology of human differences.* (3rd ed.) New York: Appleton-Century-Crofts, 1965.

U.S. Commission on Civil Rights. *Racial isolation in the public schools.* Vol. 1. Washington, D.C.: U.S. Government Printing Office, 1967.

U.S. News and World Report, Mental tests for 10 million Americans—what they show. October 17, 1966. Pp. 78-80.

Vandenberg, S.G. Contributions of twin research to psychology. *Psychol. Bull.,* 1966, **66,** 327-352.

Vandenberg, S.G. Hereditary factors in psychological variables in man, with a special emphasis on cognition. In J.S. Spuhler (Ed.), *Genetic diversity and human behavior.* Chicago: Aldine, 1967. Pp. 99-133.

Vandenberg, S.G. The nature and nurture of intelligence. In D.C. Glass (Ed.), *Genetics.* New York: The Rockefeller University Press and Russell Sage Foundation, 1968.

Vernon, P.E. Symposium on the effects of coaching and practice in intelligence tests. *Brit. J. educ. Psychol.,* 1954, **24,** 5-8.

Vernon, P.E. Environmental handicaps and intellectual development: Part II and Part III. *Brit. J. educ. Psychol.,* 1965, **35,** 1-22.

Walters, C.E. Comparative development of Negro and white infants. *J. genet. Psychol.,* 1967, **110,** 243-251.

Wheeler, L.R. A comparative study of the intelligence of East Tennessee mountain children. *J. educ. Psychol.,* 1942, **33,** 321-334.

White, S.H. Evidence for a hierarchical arrangement of learning processes. In L.P. Lipsitt & C.C. Spiker (Eds.), *Advances in child development and behavior.* Vol. 2. New York: Academic Press, 1965. Pp. 187-220.

Willerman, L., & Churchill, J.A. Intelligence and birth weight in identical twins. *Child Developm.,* 1967, **38,** 623-629.

Wilson, A.B. Educational consequences of segregation in a California community. In *Racial isolation in the public schools,* Appendices, Vol. 2 of a report by the U.S. Commission on Civil Rights. Washington, D.C.: U.S. Government Printing Office, 1967.

Wiseman, S. *Education and environment.* Manchester: Manchester Univer. Press, 1964.

Wiseman, S. Environmental and innate factors and educational attainment. In J.E. Meade & A.S. Parkes

(Eds.), *Genetic and environmental factors in human ability.* New York: Plenum Press, 1966. Pp. 64-80.

Wright, S. Statistical methods in biology. *J. Amer. stat. Ass.,* 1931, **26,** 155-163.

Young, M., & Gibson, J.B. Social mobility and fertility. In J.E. Mead & A.S. Parkes (Eds.), *Biological apsects of social problems.* Edinburgh: Oliver and boyd, 1965.

Zazzo, R. *Les jumeaux, le couple et la personne.* 2 vols., Paris: Presses Universitaires de France, 1960.

Zigler, E. Familial mental retardation: a continuing dilemma. *Science,* 1967, **155,** 292-298.

Zigler, E. The nature-nurture issue reconsidered: A discussion of Uzgiris' paper. Paper read at Conference on Sociocultural Aspects of Mental Retardation, Peabody College, Nashville, Tenn., June, 1968.

====== Here ends Professor Arthur Jensen's article. ======