The social-class, ethnic, and racial differences in performance on tests of intelligence and scholastic achievement are discussed. Significant deficiencies exist in the means of lower-class and certain racial groups for many measures of ability, motivation, and performance. Composite attempts to estimate the proportion of phenotypic variance in IQ which is genetic approximate 80%. These heritability indices have been interpreted to mean that most of the observed deficiencies in the mean IQ's for classes and races is biologically inevitable. However, evidences of plasticity suggest that the range of reaction in the IQ must be of the order of 75 points. Such evidence of plasticity becomes less puzzling when one recognizes that estimates of heritability indicate the loss of deviation from the mean of the population to be expected from parents to off-spring in experiments on selective breeding, that indices of heritability are relevant only to the status quo within a given population so long as environment remains constant, and that these indices say nothing about how much the mean of a phenotypic measure of intelligence or scholastic achievement will change with development in new environments. Such information demands knowledge of the norm of reaction which comes only from the difference between the means of measures of achievement and intelligence for groups of children from a given population of genotypes who have been reared under differing environmental conditions or differing educational programs. Evidence from two sources indicates that this range is of the order of 75 points and that special child rearing can boost the mean achievement for disadvantaged children. (DB)
HEREDITY, ENVIRONMENT, AND CLASS OR ETNIC DIFFERENCES

by

J. McV. Hunt

University of Illinois

1. This paper was prepared at the invitation of the Chairman, Professor Anne Anastasi of Fordham University, for the 1972 Invitational Conference on Testing Problems sponsored by the Educational Testing Service. The preparation was supported by grants from the U.S. Public Health Service: MH-K6-18567, MH-11321, and MH-16074.
HEREDITY, ENVIRONMENT, AND CLASS OR ETHNIC DIFFERENCES

by

J. McV. Hunt

University of Illinois

That people differ on the average according to the social classes and the ethnic or racial groups to which they belong is an ancient observation. Within our own United States, the mean values or the incidence of behavioral phenomena have been observed to vary significantly for nearly every characteristic where systematic measurement has been tried. Such characteristics include emotional adjustment as measured by both the Minnesota Multiphasic Personality Inventory (Gough, 1946, 1948) and the Vineland Social Maturity Scale (Sims, 1954), the incidence of behavior disorders (Clark, 1949; Faris & Dunham, 1939; Landis & Page, 1938), the incidence of delinquent and criminal behavior (Shaw, 1929; Short & Strodtbeck, 1965), persistence in goal striving (Battle & Rotter, 1963; Hertzig, et al., 1968; Zigler & Butterfield, 1968), the tendency to be impulsive rather than reflective (Meichenbaum & Goodman, 1969; Mumbauer & Miller, 1970), measures of ability, and measures of both scholastic and occupational achievement. Such differences serve to define the social-class structure as described by Warner and his colleagues (see Warner, et al., 1949). Differences among ethnic and racial groups include a similar variety of characteristics plus differences of language, skin color, and cultural attitude.

Especially prominent have been the social-class, ethnic, and racial differences in performance on tests of intelligence and scholastic achievement. Since the evidence is most abundant for such measures, they shall be the concern of this essay. Where social class is the focus, the IQs of children entering school from professional families average about 115 while those from unskilled laborers average about 95 (McNemar, 1942), and differences in mean IQ of this order of 20 points are typical of those reported by other investigators (see Anastasi, 1958). Where race is the focus, Black children usually average an IQ about one standard deviation, or 15 points below white children (Shuey, 1966). Moreover, Mayeske (1971) has described a special analysis of the data in the report on equality of educational opportunity by Coleman, et al., (1966) which shows that 25 percent of the total variance among American students in their academic achievement was associated with membership in one of six ethnic-racial groups: Indian, Mexican, native white, Negro, Oriental, and Puerto Rican.

Although some individuals have risen out of poverty to top levels of excellence, there can be no blinking these class and race differences. They exist. But are they biologically inevitable to the degree in which they now manifest themselves? Are class differences an inevitable matter of competitive social selection which has resulted from genotypic limits on potential? Or, might one expect most of such differences as exist from the conditions of child-rearing in the various social classes? Are the observed ethnic and racial differences merely a matter of the relative frequency in which certain genes are distributed in these ethnic groups? Or, inasmuch as the predominant majority of Indians, Mexicans,
Negroes, and Puerto Ricans have lived since birth in conditions of poverty in families of little education and little hope of advancing status, might one justifiably expect what they show on the tests and in the schools from the conditions of their rearing?

Such questions are usually taken to involve the old issue of nature and nurture and to pose the question of their relative potency. Let me begin by recognizing that heredity is clearly primary. For such a matter as whether a given fertilized ovum becomes a human being or a member of another species, heredity is all important. One gets only elephants from breeding elephants, nothing else. Moreover, each individual begins life with a given complement of genes which he received in equal shares from each parent. The DNA in his genes carries information which exerts a continuing influence on his development throughout his life. Since a life begins with conception, heredity is always primary. Yet, having such a primary status and exerting a continuing influence throughout life need imply neither genetically fixed traits nor a predetermined course and rate for later development.

WAYS OF OBTAINING ANSWERS

One approach to answering such questions comes by way of definition. Cyril Burt (1969) has recently contended that intelligence should be defined as "innate, general, cognitive ability." Such a contention has been the dominant view in England ever since Sir Francis Galton (1869), Charles Darwin's younger cousin, conceived genetically fixed traits as implicit in the theory of natural selection, launched the study of individual differences with the publication of Hereditary Genius, and invented the correlation statistic to show the persistence of individual differences across generations. Karl Pearson, Galton's successor, improved on the correlation method, extended the scope of such investigation, and used the evidence to support his own such definition (Pearson, 1902, 1904).

Defining intelligence as innate ability has the defect of leaving it unmeasurable. Or, if scores on tests of intelligence are taken as measures of innate ability, this approach confuses an observable and measurable phenotype with the genotype which can only be inferred -- a distinction which Johannsen (1903, 1909), one of the fathers of scientific genetics, made at the turn of the century. Ignorance or neglect of this distinction, has all too often left psychologists discussing measures of intelligence as if they were genotypic. Moreover, the traditional distinction between tests of intelligence or aptitude from achievement tests tends to maintain such confusion. Here it has been the merit of Lloyd Humphreys (1962) to recognize that these two kinds of tests differ only in degree with intelligence tests including a wider variety of items, tapping a wider variety of experience, and being further removed in time from relevant learning situations than achievement tests.

An approach by definition can start quite as justifiably with the contention that intelligence is primarily a product of learning rather than its cause. Thus, George Ferguson (1954, 1956) has explained abilities as derived from factor analysis as the results of transfer of training in overlearned skills. Moreover,
Gagné (1968) has viewed mental ability as a product of cumulative learning in which various skills form a transfer hierarchy ranging from stimulus-response connections through chains, motor and verbal, multiple discriminations, concepts, and rules, simple and complex. In the recognition of a hierarchical structure, Gagné's view resembles that of Piaget (1937, 1947), and my own (Hunt, 1961), yet Gagné's view gives less importance than either Piaget's or mine to the effect of the individual's existing organization of abilities on the consequences of his encounters with specified circumstances. Clearly, a definitional approach, by itself, can lead only to fruitless debate. The scores of all kinds of tests of abilities represent past developmental achievements in which the influence of the genes have interacted continuously with the influence of the environment in an on-going series of encounters.

The second approach is empirical. It is based upon the correlations between test scores from relatives and the operations are quite independent of these belief-based definitions. Such correlations have been found regularly to be a positive function of the degree of genetic relatedness. Thus the correlations between monozygotic or identical twins are higher than those between dizygotic twins or siblings, or those between parents and children, which, in turn, are higher than those between cousins, and those between first cousins are somewhat higher than those between second cousins or unrelated children reared apart (see Erlenmeyer-Kimling & Jervik, 1963; Fuller & Thompson, 1960, Chap. 7; Outhit, 1933). Moreover, the correlation between the scores for siblings reared apart \( r = +.34 \) in Freeman, Holzinger, & Mitchell, 1928; median \( r = +.47 \) from 33 studies in Jensen, 1969, p. 49) is higher than that reported for first cousins \( r = +.06 \) in Gray & Moshinsky, 1933; median \( r = +.26 \) from 3 studies in Jensen, 1969). Such findings and others (see Jensen, 1969) clearly attest a genetic influence on phenotypic measures of intelligence.

It has been more difficult to say how great this influence is. The traditional effort to separate and specify the relative strengths of the influences of heredity and environment has led to a variety of research and statistical designs beyond the scope of this essay (see Cattell, 1953, 1960; Falconer, 1960; Fuller & Thompson, 1960). The answer to the question has come in the form of estimates of heritability. Heritability is defined as that proportion of the variance within a specific population in the phenotypic measure of a characteristic that is determined by the genotypic variation within that population (Rieger, Michaelis, & Green, 1968, p. 212). In selection experiments, a simple approximation of heritability may be obtained by dividing the gain in the offspring by the selection differential. This gain is the difference of the mean of the trait measures concerned for the offspring from the mean for the population. The selection differential is the difference of the mean for those selected to be the parents from the mean of the population. The closer this ratio is to unity, the higher the heritability.

Since experimental selective manipulation of human matings has been out of the question for many reasons, estimates of heritability for human intelligence have been based upon statistical manipulations of the correlations between relatives. The findings have varied (see reviews by Erlenmeyer-Kimling & Jervik, 1963;
Galton (1883) began this effort with his studies of twins and concluded that "nature prevails enormously over nurture when the differences of nurture do not exceed what is commonly to be found among persons of the same rank of society and in the same county" (p. 241) although he lacked a numerical estimate. Such an estimate based upon the correlation between identical twins reared apart is conceptually the most obvious, if their environments are uncorrelated, because they have only their genes in common. Newman, Freeman, and Holzinger (1937) reported a correlation of +.77 between the Stanford-Binet IQs of their 19 pairs. More recently, according to Jensen (1969), Shields (1962) has also reported a correlation of +.77 between composite scores from 44 pairs on Raven's Progressive Matrices, and most recently, Cyril Burt (1966) has reported a correlation of +.86 between the Stanford-Binet IQs of 53 pairs of identical English twins. Such findings support both Galton's statement and the oft-quoted numerical estimate that 80 percent of phenotypic variance in the IQ derives from heredity.

Other bases for estimating heritability yield estimates differing little from the correlation between measures of intelligence for identical twins. The median of the correlations between such measures in five samples of unrelated children reared together is +.24. Since all these have in common is their environments, this serves as an estimate of variance in phenotypic intelligence due to environment. Subtracted from 100%, it leaves 76% due to heredity. Another basis comes from an attempt to assess the extent of family resemblance with cultural differences held constant by means of what was contended to be a culture-free test. Again, approximately 80% of the variance in phenotypic measures of intelligence among families was attributed to heredity (Cattell & Willson, 1938). When Jensen (1967) applied his generalized formula for estimating heritability from any two correlation sets of relatives where the degree of kinship is higher for one than the other to all those reported, he got a composite heritability value of +.77 (+.81 when corrected for unreliability). This he regards as the best overall estimate of the heritability of human intelligence (Jensen, 1969, p. 43). (Just how any single estimate can be expected for such a population statistic which varies with the variability of both genetic and environmental factors, he does not explain.)

Before World War II, textbooks of psychology often stated that heredity accounts for approximately 80% of the variance in the IQ. This is the assertion and the line of reasoning that I characterized, largely for purposes of exposition, as "the belief in fixed intelligence" (Hunt, 1961). Along with this belief went another that behavior unfolds directly with genetically controlled maturation which I termed "the belief in predetermined development" and traced back to G. Stanley Hall and his elaboration of what he saw as the implications of the recapitulation doctrine. This latter belief set the conceptual stage for the claim of a constant IQ and the notion that individual differences start with conception and remain essentially constant throughout development and life. It is basically this same argument which Arthur Jensen (1969) has revisited to explain what has come to be characterized as "the failure of Project Head Start." Implied also is the proposition that if 80% of the variance in phenotypic measures of intelligence derive from heredity, then most of those differences observed between the means of the IQs for children of professional and unskilled parents or for the children of parents in the racial groupings must be biologically inevitable.
THE DISSONANT EVIDENCE OF DEVELOPMENTAL PLASTICITY

This claim that 80% of the variance in phenotypic measures of intelligence and developmental advancement are attributable to heredity makes other evidence of developmental plasticity highly puzzling. Bits of such evidence began turning up even before World War II. They included the increases in the IQs of preschool children associated with nursery schooling in the Iowa studies (Skeels, Updegraph, et al., 1938; Wellman, 1940), lack of longitudinal predictive value for scores from tests given infants during their first two years even though these scores exhibited satisfactory reliability (Bayley, 1940), and the finding by Skeels and Dye (1939) of dramatic increases in the IQs of orphanage-reared infants who were moved to an institution for the mentally retarded where "the older and brighter girls on the ward became very much attached to the children and would play with them during most of their waking hours" (p. 5, see Hunt, 1961). Instead of calling the beliefs in "fixed intelligence" and "predetermined development" into question, such was the firmness of faith in them that such evidences merely evoked a flood of methodological criticisms tending to discredit them.

More of such evidence has come since World War II. René Spitz (1945, 1946a, 1946b) and Wayne Dennis (1960) both reported apathy and dramatic retardation, even in locomotion, associated with orphanage rearing. Students of neonatal behavior turned up unsuspected evidences of ability for both classical and operant conditioning during the first few days following birth (for reviews see Lipsitt, 1963, 1966, 1967). Investigations stemming from the neuropsychological theorizing of Donald Hebb (1949) brought evidence that increasing the complexity of early perceptual experiences of rats increases substantially their later ability to learn the Hebb-Williams (1946) mazes (Forgays & Forgays, 1952; Forgus, 1954; Hymovitch, 1952). According to this same principle, both rats (Hebb, 1947) and dogs (Thompson & Heron, 1954) which were pet-reared became significantly superior as adults to their cage-reared litter-mates in learning these mazes. For many skeptics, such evidences from the animal laboratory obviated selectivity and regression toward the mean, two of the criticisms commonly leveled against the evidence from the orphanage and nursery-school studies, as the basis for such early effects of experience on development. Thus, they lent also a cubit of credibility to the orphanage evidence.

Other evidence has suggested that the use of a sensorimotor organization solidifies its structure and hastens its development. In one such study, providing infants, beginning at 5 weeks of age, with a stable pattern over their cribs to look at increased the age at which the blink response appeared, to target drops of 11.5 inches, from 10.4 weeks of age for 10 control infants without such an opportunity to use their eyes to an average of 7 weeks for 10 infants provided with stable patterns (Greenberg, Uzgiris, & Hunt, 1968). In another such study, providing infants with visual targets of complexity properly matched to their level of development increased the age at which they achieved mature reaching for a seen target from a median of 145 days, for infants in an original normative study, to a median of 89 days for the infants in a second enrichment study where the complexity of the visual stable was properly matched to the development of the children (White, 1967). If one casts these findings into the terms of Stern's (1912) IQ ratio in order to put them in familiar perspective, lowering the age for the
appearance of the blink-response from 10.4 weeks to 7 weeks is a gain of approximately 48 points, and lowering the achievement of that visuo-motor coordination in mature reaching from a median of 145 days to a median of 89 days is a gain in the order of 63 points. Such a transformation applies only to past developmental achievement. It should imply no permanence unless the circumstances of these infants are so arranged as to give them special opportunities to accommodate their advanced visuo-motor skills to new situations calling for further development.

Evidence of plasticity comes also for abilities more closely akin to tested intelligence and scholastic proficiency in what Piaget (1936, 1937) has termed object construction and imitation. We have developed a set of six sequentially ordinal scales (Uzgiris & Hunt, 1973). One is for object permanence and involves what is probably the most basic epistemological construction, and the other to be discussed here is for the development of vocal imitation.

Such ordinal scales enable one to define a level within a line of development existing between the top landmark passed and the first one failed. In cross-sectional studies, one can then compare the ages of children at given levels of development who are living under differing environmental conditions. In Athens are two orphanages for illegitimate babies, a municipal orphanage where the infant-caretaker ratio approximates 10/1 and Metera, which attempts to be a model baby center, where this ratio is of the order of 3/1. Whether an illegitimate infant comes to one or the other of these orphanages appears to be a matter of chance. We have compared the ages of all the children from these two orphanages who were at the several levels on these scales of object permanence and vocal imitation; we also included a sample of children from working-class families. The results for object permanence appear in Figure 1. For illustrative purposes consider the means and the standard deviations for that level of object permanence where a child follows an object through one hidden displacement but not through a succession of them. These appear in the left three of the middle cluster of vertical bars. For the Municipal Orphanage, the tallest bar on the left, the mean age for the five children at this level is 143.76 weeks with a S.D. of 29.19 weeks; for Metera the mean, also for five children, is 94.39 weeks with a S.D. of 11.13 weeks, and for the home-reared children, third bar, the mean age for the 16 children at this level is 87.9 weeks with a S.D. of 28.06. The children of the Municipal Orphanage average significantly older than those from either Metera or the working-class homes who differ not significantly.

Consider also that level of object permanence at which a child follows an object through a series of hidden displacements and reverses the order in which the container of the desired object disappeared but fails to copy a series of five sticks decreasing in length from bottom to top. This is the top level on our scale of object permanence. For the Municipal Orphanage, the mean age for 7 children at this level is 206.58 weeks with a S.D. of 34.29; for Metera, the mean age of 4 children is 154.58 weeks with a S.D. of 17.32 weeks; and for the 16 home-reared children at this level the mean age is 130.77 with a S.D. of 47.11 weeks. Again, the children of the Municipal Orphanage are significantly older than either of the other groups, and these do not differ significantly. It also worth noting the S.D.s. The smallest is that for Metera where the conditions of rearing are most nearly alike, next for the Municipal Orphanage where the child-
OBJECT CONSTRUCTION UNDER DIFFERING CONDITIONS OF REARING

Child searches for and obtains desired object when it is:

- Completely covered
- Covered by 3 screens
- One hidden displacement
- Series of hidden displacements
- Order of container disappearance
- Reversal of order of container disappearance

Method
- Cross-sectional single examination
- Longitudinal repeated examination

Mt. Carmel, Ill.

MOS Orphanage - Infant/caretaker ratio: about 10:1
MOS Orphanage - Infant/caretaker ratio: about 3:1
House raised

from Pacansky & Hunt (1977)
Uzgiris (in press)
Hunt, Bachrach, Badger & Hoo

Figure 1
caretaker ratio of 10/1 almost insures the choice of pets while others are neglected, but the S.D. is largest by about 13 weeks for the home-reared children (Paraskevopoulos & Hunt, 1971). Presumably variations within families combine with differences in heredity to exaggerate the variance in the development of object permanence.

More directly relevant to the issue of the degree to which the class differences commonly observed are biologically inevitable is the evidence from two longitudinal studies made for quite other purposes. In one, a still unpublished study by Ina Uzgiris, the scales were administered every other week to 12 infants from middle-class homes. The other, also a still unpublished study by Schickedanz and myself, was made to evaluate the effects of a mothers' training program (Badger, 1971, 1972) on the development of the children of poverty being reared by parents participating in the Parent-and-Child Center of Mt. Carmel, Illinois. Here again, the scales were administered every other week to 8 consecutive infants of these parents from the poverty sector who were participating in the program of this Center. Let us consider the means and standard deviations of the ages at which these two samples of children achieved the same two levels of object permanence. In Figure 1, the results for these two samples appear in the two bars on the right of each cluster. For following a desired object through one hidden displacement, note again the middle cluster. The mean age for the children of middle-class homes was 58.46 weeks (S.D. = 7.43 weeks) and that of the infants of parents of poverty who were participating in the mothers' training program was 48.06 weeks (S.D. = 9.22 weeks). All eight had achieved this level by one year of age, and the youngest by 44 weeks of age. For the top level of object permanence, following a desired object through a series of three hidden displacements in a reverse order, the mean age for those of middle-class was 91.36 weeks (S.D. = 9.43) and for the children of poverty in the Parent-and-Child Center was 72.74 weeks (S.D. = 13.99). Apparently this mothers' training program has hastened the development of object permanence. It is often assumed that the child-rearing of middle-class families approximates the optimum, but here is an instance in which the child-rearing of the parent caretakers from the poverty sector has been improved by a program of mother training to surpass that of the middle-class -- at least for object permanence during the first two years of infancy. If support for the Parent-and-Child Center program is continued long enough, we shall be able to follow the development of these children and also their performance into school to see how their performance compares with that of their older siblings who lived through their 'early years before their mothers were touched by Mrs. Badger's program of mother-training. We desperately need the kind of evidence one can only get from such prolonged longitudinal studies of children developing under differing environmental circumstances.

Only class differences have been concerned in these comparisons for object permanence. All of the children in both samples were white. But evidence of plasticity in phenotypic measures of intelligence comes also for Black children from a demonstration underway in Milwaukee which is directed by Heber and Garber (1972), of the University of Wisconsin. According to a progress report (October, 1970) to the agency supporting their work, Heber and Garber started their investigation with a survey of tested intelligence in that census tract of Milwaukee with the lowest socio-economic status. From this survey came the interesting finding that approximately 80% of those children with IQs below 80 came from families where the mothers had IQs below 80. This finding fit well the hypothesis
of biologically inevitable differences, but these investigators did not stop at this point. Rather, they selected a sample of 40 high-risk mothers with new infants. These mothers were identified objectively by having IQs of 75 or below. For 20 of these families, a home visitor saw and played with the infants until each infant was approximately 6 months old. Heber and Garber then arranged to have the infant brought 5 days a week to a day-care center. There each was cared for by a woman who had been selected for articulate speech and who had been trained to provide appropriate educational experiences for infants. Each infant was also put on a program of repeated testing, and the mother was given counseling in homemaking and in the buying and preparation of food. For the other 20 families, the program was limited to routine counseling visits with the mothers and what I believe was a duplication of the program of repeated testing for their infants. At age 45 months, the IQs of the children of the control families average approximately 90. This is a whole S.D. above that for their mothers, an unusual degree of increase which probably derives in part from the repeated testing as well as from the expected regression effect. At this same age of 45 months, the IQs of those infants provided with educational day-care are reported to average an almost unbelievable 128. Unless there is something very wrong with this demonstration that I cannot now see, it provides spectacular evidence of plasticity in a standard phenotypic measure of intelligence within Black children from families of the highest risk where mothers have IQs near the low end of the scale. Professor Heber has been properly cautious about attributing permanence to such a gain.

Moreover, I suspect considerable loss is inevitable if these children are simply returned to their families and to their neighborhood schools. Tests of intelligence measure only past achievements. They say very little of the future unless the circumstances of future development are specified.

Maturation and Experience

In our traditional conception of development, maturation and learning have represented completely separate processes. Maturation has been conceived to be controlled by the genes. Learning has been conceived to be controlled by environmental encounters. Clear evidence has come, however, that informational interaction through the eyes influences maturation within the central nervous system. For the most part, these investigations have been inspired by the neuropsychological theorizing of Donald Hebb (1949) and the neurobiochemical theorizing of Helgar Hyden (see 1959). Apparently inspired by Hebb's theorizing, Austin Riesen reared chimpanzees in the dark. The dark-rearing resulted not only in behavioral deficits but also in a diminution of the number of nerve cells and glial fibers developing within their retinal ganglia by adulthood, (Chow, Riesen, & Newell, 1957). Corroboratively, Brattgärd (1952), inspired by Hyden's biochemical theory of memory, reported that rearing rabbits in the dark results in a paucity of RNA production in their retinal ganglia as adults. Since then, a California group (Bennett, Diamond, Krech, & Rosenzweig, 1964; Krech, Rosenzweig, & Bennett, 1966) has reported that thickness of the cerebral cortex and the level of total acetylcholinesterase activity of the cortex, as well as rate of adult maze-learning, are the function of the complexity of the environment during early life. Quite recently, studies of the effects of dark-rearing during early life have been extended through the visual system. Wiesel and Hubel (1963), for instance, have demonstrated that dark-rearing produced a paucity of both cells and glial fibers in the lateral
geniculate body of the thalamus, and a Spanish investigator by the name of Valverde, and his collaborators, have obtained evidence that dark-rearing also decreases both dendritic branching and the number of spines which develop on dendritic processes of the large apical cells of the striate area in the occipital lobes in mice (Valverde, 1967, 1968; Valverde & Esteban, 1968). Evidence that dark-rearing diminishes higher-order dendritic branching in cats as well as mice has been reported by Coleman and Riesen (1968). Evidence suggesting that such effects on dendritic branching and spine density may be a matter of the complexity of the information encountered and the variety of adaptations called forth has come from studies by Holloway (1966) and Schapiro and Vukovich (1970). Most recently, Fred Volkmar and William Greenough (1972) at the University of Illinois have compared the dendritic branching of stellate and pyramidal neurones, Golgi stained, in several layers of the occipital cortex for litter-mate rats reared in a complex environment (where a group of 12 pups were housed in a large wire-mesh cage provided with a variety of toys that were changed daily), in environments consisting of pairs of animals in standard laboratory cages or single animals in such cages. Those reared in the complex environment, comprised of an enclosed area of several square feet filled with manipulable objects changed regularly, exhibited considerably greater branching of dendrites of the third and higher orders. They showed seven times as many fifth-order branches on the pyramidal cells of the Layer V as their litter-mates reared in isolation and about 3.5 times as many as their litter-mates reared in social pairs. Since light of about the same intensity was available for all, it would appear to be complexity of experience rather than mere absence or presence of light which is responsible for these very substantial differences. Not only do such findings show a great deal of environmental plasticity in neuroanatomical maturation; they also suggest that the variety of informational inputs from circumstances of greater complexity call forth accommodative adaptations which show in neuroanatomy as well as in behavior. If one considers 80% of the variance in phenotypic measures of intelligence and related matters to be relevant to such evidences of plasticity in behavior and neuroanatomical maturation, then all this is highly puzzling, especially so since a partitioning of the variance in this study showed more than half of it related to environmental conditions (Volkmar & Greenough, personal communication).

THE NORM OF REACTION

Let us consider how much variations in environmental conditions appear from existing evidences to be able to alter phenotypic measures of intelligence.

Ever since Walterec introduced the concept in 1909 (see Dunn, 1965), geneticists have concerned themselves with the "norm of reaction" or the "range of reaction" as well as with Mendelian statistics and the mechanisms of genetic transmission. The "norm of reaction" is defined as the range of phenotypic reactions which a specified genotype is able to produce in response to environmental influences (Rieger, Michaelis, & Green, 1968, p. 372). Such a concept, like that of heritability, can never be fully specified from empirical data because a new investigator with imagination can always arrange a new progression of environmental encounters which may alter further the range of phenotypic reactions. In investigative practice, however, one obtains relevant evidence in terms of the difference between the means
of phenotypic measures of a given trait for samples of individuals from a given population of genotypes who are reared in different environments. For measures of intelligence and scholastic ability, one can get evidence from comparing the mean values of phenotypic measures for individuals from a given population reared under differing environmental circumstances or differing educational programs.

Such data are still very few. In investigations with human subjects, moreover, they are seldom based strictly on a single population of genotypes; nevertheless, highly suggestive evidence exists, and one set comes from the ages at which children achieve the various levels of object permanence. Here the Athenian children in the study by Paraskevopoulos and Hunt (1971) comprise about as close an approximation of a single population of genotypes as one can expect to get. They represent the lower half of the socio-economic structure. The American children in the still unpublished Worcester study by Uzgiris represent the middle-class, but those from the Parent-and-Child Center at Mt. Carmel by Schickedanz and Hunt represent families of the lowest socio-economic status who were recruited from those on Welfare and on Aid to Families with Dependent Children. For following a desired object through a series of hidden displacements in reverse order, the extreme mean ages are 206.58 weeks, for the children of the Municipal Orphanage where the child-caretaker is 10/1, and 72.74 weeks for the children of poor families reared with the aid of the Badger Mother's Training Program in the Parent-and-Child Center at Mt. Carmel. For the Worcester children from middle-class families, the mean is 91.36 weeks with S.D. of 9.43 weeks. Even though I have been developing ordinal scales in part to escape Stern's IQ ratio and the normative approach to a meaning for test performances and scores (Hunt & Kirk, 1971), it may be worthwhile here for purposes of communication to cast these figures into this familiar ratio by assuming that 91.36 weeks for the children of middle class approximates the norm of 100. Thus, this empirical evidence indicates that the range of reaction must extend at least from a low of about 50 to a high of about 125, a range of reaction of 75 points of IQ.¹

One may well object that such a range could be found for only a simple function during infancy when the longitudinal validity of measures of intellectual development is low. But this empirical difference of 75 points is essentially the same as that found by Wayne Dennis (1966) when he get the Goodenough Draw-A-Man Test given to samples of typical children, aged between 6 and 9 years, who were living in typical family environments in some 50 cultures over the world. The variation in the means of such IQs ranged from a low of 52 to a high of 124, a range of 72 IQ points. The variation in this phenotypic measure appears to be associated with the degree of contact with and participation in representative graphic art. Probably this Draw-A-Man Test calls for a considerably less complex set of abilities, as these are assessed by factor-analysis, than either of the

¹ The lower limit of this range deserves a word of explanation. Dividing 91.36 weeks by 206.58 weeks yields less than .5, but since the mean of 206.58 weeks derives from a cross sectional study, it must exaggerate the delay more than would a longitudinal approach with examinations every other week. For this reason, I have rounded the lower limit to the approximation of an IQ of 50.
more standard scales: the Stanford-Binet test or the Wechsler-Bellvue Children’s Scale. Yet, for American children, the IQs from the Draw-A-Man Test correspond about as well with IQs from these more standard tests as do IQs from either of these other two tests with each other. It must also be admitted that children in a cross-cultural situation cannot come from a single population of genotypes, yet typical individuals from the Syrian nomadic tribe have shown the capacity to adapt themselves readily to technological cultures when reared in them.

These empirically determined ranges of 72 and 75 points in mean IQs fall only about one S.D. short of the full range of individual IQs (between 55 and 145) which includes all but a fraction of a percent of individuals above the pathological group which bulges at the low end of the distribution for the IQ. Clearly, there is dissonance between any argument based on the statistics of heritability and this argument from ranges of reaction.

RELEVANCE: HERITABILITY VERSUS NORM OF REACTION

Perhaps this dissonance can be clarified by the respective meanings of heritability and the norm of reaction. Heritability is, by definition, that proportion of the total phenotypic variance for a particular characteristic in a specified population (Rieger, Michaelis, & Green, 1968, p. 212). Heritability is not an attribute of a trait, but rather of a trait in a specified population developing and living within a given set of environmental conditions. What a given estimate of heritability gives is the amount of gain or loss to be expected in the course of selective breeding. Thus, given heritability for a particular trait in a given population at 80%, if a sample of parents is selected to have a mean parent measure of the trait averaging one S.D. above (or below) the mean for the population, then the mean of the measures for that trait in the offspring would be expected to average .8 of a S.D. above (or below) the mean of the population. This expectation would hold, however, if and only if the environmental circumstances remained constant through the lives of the two generations. Thus, an index of heritability tells us about how much of the selection advantage or disadvantage is lost between parents and their offspring.

On the other hand, an index of heritability cannot tell us nothing about how much change in the measures of a phenotypic trait will result from being reared in new environments. It can tell us nothing about how much the IQs of the children from a given population of genotypes will be changed through being reared in newly designed environments and educational programs. Thus, a composite heritability index of 80% for the IQ may say how much of the variance in IQ is hereditary for the kinds of children studied who have developed under the existing conditions.

2 This dissonance between the evidences for a composite index of the order of 75% or 80% for heritability of the IQ and the evidences of plasticity in development has long puzzled me. I am greatly indebted to writings of Jerry Hirsch (1970, 1972) and to discussions with him for this clarifying interpretation of this dissonance.
environmental conditions of American and European culture, yet it tells us nothing of how much the IQ might be changed by newly designed systems of child-rearing and education. It is not relevant to why Project Headstart succeeded or failed. Knowledge of how much the IQ can be altered by new regimes of child-rearing and early education can be obtained only from evidences of the range of reaction or from studies of the differences between the means of IQ for children reared in differing environmental conditions and with differing programs of early education.

In a dynamic and developing society, the conditions of child-rearing and education are always changing, hopefully improving. Measures of phenotypic intelligence would be expected to go up with these supposed improvements unless the increases in intelligence are hidden in comparative scores based on new norms. They do. In a number of repeat-studies, increases in average IQ have occurred rather than the loss predicted by Cattell (1937) from differential fertility (see Hunt, 1961). Cattell (1950) himself published one of these based on test surveys of the children in an English city in 1936 and 1939. Instead of the predicted drop of one point, he found a gain in mean IQ of 2.28 points. Others have reported substantially larger gains in mean IQ. One by Smith (1942), based on surveys of the children aged 10 to 15 in the schools of Honolulu in 1924 and in 1933, reported a gain in mean IQ of 20 points. Another by Wheeler (1942), based on tests, made 10 years apart, of samples of students aged 10 to 12 years from given families in the schools of the Tennessee Valley before and after the changes introduced by the Tennessee Valley Authority, reported a 10-point gain in median IQ. Another by Finch (1946), based on tests given to all of the children in a sample of high schools in Minnesota during the 1920s and again in the 1940s, reported gains in mean IQs for the various high schools ranging from 10 to 15 points. In yet another such study, Tuddenham (1948) compared a representative sample of draftees from World War II with a sample from World War I on comparable forms of the Army-Alpha Group-Test. The median for World War II fell at the 82nd percentile of the distribution of World War I. Thus, half of the draftees of World War II belong among the upper fifth for World War I (see Hunt, 1961, p. 337ff). So long as one considers a high index of heritability relevant to educability, such gains would seem incredible, but indices of heritability are not relevant.

IMPLICATIONS FOR CLASS AND RACE DIFFERENCES

The implications of these considerations for class and race differences become readily apparent when one considers the inequalities of environmental opportunity across the class structure of American and European societies. Attempts to assess the genotypic potential behind phenotypic measures of intelligence have always assumed essentially equal environmental opportunity for growth, adaptation, and learning with micro-inequalities randomly distributed. The past two decades, however, have yielded abundant evidence of large deficiencies in the development-fostering quality of the environments provided for children of lower-class families of poverty whatever their ethnic origin and race. Inasmuch as a major share of Indian, Mexican, Puerto Rican, and Black families fall in the poverty sector, their children share the poverty-based deficiencies of poor white families plus whatever additional disadvantages are associated with dark skins and differences in language.
The deficiencies in the development-fostering quality of the environments associated with poverty I have reviewed elsewhere in some detail (Hunt, 1969, pp. 202-214). They include basic nutritional deficiencies in a substantial share of mothers at the time of conception and during gestation. They include a lack of opportunities to acquire cognitive and linguistic skills illustrated by such facts as the following: that where approximately 90% of nursery school children can respond by picking up appropriate blocks for the colors named by an examiner and approximately 80% can name all of the six colors used, only approximately 20% of 4-year-olds beginning a Headstart program can respond in these fashions (Kirk & Hunt, in preparation). These deficiencies also include a lack of opportunities to develop the motivational systems required for confidence and persistent striving and also the opportunities to acquire those values and standards of conduct demanded by the mainstream of a complex organized society. These are not small variations in environmental opportunity, and certainly they are not randomly distributed across the class structure.

Given the evidences of plasticity indicative of a range of reaction for measures of phenotypic intelligence of the order of 75 points, and given these class and race differences associated with poverty in the development-fostering quality of the environments, and especially of the early environments, provided for children, relatively small portions of the commonly reported deficits in the means of IQ and measures of scholastic performance for the children of unskilled white parents, Black parents, Indians, Mexicans, and Puerto Ricans can be considered to be biologically inevitable. To be sure, the evidence and the argument summarized does not rule out a contribution from heredity to these differences. Yet if the Mother's Training Program of Earla Deen Badger can bring the average IQ of even a small sample of families of poverty approximately 25 points above that for middle-class families (as assessed by the scale of object permanence), and if the Heber-Garber program can bring the average Stanford-Binet IQ for children of Black mothers with IQs of 75 or below up to 128 at age 45 months, it becomes hard to believe that more than a very minor share of the differences among class and ethnic groups are biologically inevitable.

Recently, this case against the biological inevitability of class and race differences has received empirical support from another direction. In a study reported at the Washington meetings of the American Psychological Association, George W. Mayesek (1971) reported a special analysis of the data in the report on Equality of Educational Opportunity by Coleman, et al. (1966) designed to ascertain the degree to which that 25% of the variance in scholastic achievement associated with racial and ethnic group membership could be explained in terms of socio-economic and educational circumstances. From the relevant partial regression equations, he took into account the socio-economic status of each family, the presence or absence of key members of each family, the assessments of the aspirations for schooling by students and parents, parental beliefs about how students might benefit from education, their region and neighborhood of residence, and the achievement and motivations of the students attending the school. When this was done, the variance among the students in their academic achievement scores associated with ethnic-group membership dropped to 1.2%. Similar findings have been reported by Jane Mercer (1971) for Chicano and Black children in the schools of Riverside, California. The more the families of these children
resembled those of the modal configuration for the middle-class, white community of Riverside in terms of five characteristics, the more nearly did the mean IQs of the children approximate 100. In these studies, both the partial regression equations used by Mayeske and Mercer's modal characteristics of middle-class families contain potentially genetic variance, but the force of such a consideration is reduced by the evidence of IQs well above average for children of poor white families and of Black mothers with low IQs when those children are provided with experiences which foster their psychological development. One would guess from such combinations of evidence that all but a very minor share of children of poverty or of unfavored ethnic and racial groups could be reared in a fashion which would permit them to perform quite adequately in our technological culture if the economy provided the opportunity. Moreover, many of those now typically fated for relative incompetence might well with more fortunate rearing achieve excellence along one of the diverse avenues of achievement in our society.

RECAPITULATION AND CHALLENGE

Let me recapitulate and then present what I see as the challenge. Significant deficiencies exist in the means of lower-class and certain racial groups for many measures of ability, motivation, and performance. Most of the evidence, however, concerns measures of intelligence and scholastic performance. Composite attempts to estimate the proportion of phenotypic variance in IQ which is genetic approximate 80%. These heritability indices have been interpreted to mean that most of the observed deficiencies in the mean IQs for classes and races is biologically inevitable. These interpretations are in puzzling dissonance with evidences of plasticity which suggest that the range of reaction in the IQ must be of the order of 75 points. Such evidence of plasticity becomes less puzzling when one recognizes that estimates of heritability indicate the loss of deviation from the mean of the population to be expected from parents to off-spring in experiments on selective breeding, that indices of heritability are relevant only to the status quo within a given population so long as environment remains constant, and that these indices say nothing about how much the mean of a phenotypic measure of intelligence or scholastic achievement will change with development in new environments. Such information demands knowledge of the norm of reaction which comes only from the difference between the means of measures of achievement and intelligence for groups of children from a given population of genotypes who have been reared under differing environmental circumstances or differing educational programs. Even though such evidence is sparse, that from two sources indicates that this range is of the order of 75 points and that special child rearing can boost the mean achievement for white children of poverty and for Black children from mothers with IQs of 75 or below well above the population average.

If measures of heritability say nothing about educability, then the measures we have are irrelevant to the outcome of Headstart. But several factors elaborated elsewhere (Hunt, 1969, Chapt. 5) help to explain why Headstart is said to have failed. First, the goals were unrealistic, and, in terms of a broader view of social change, Headstart appears to have had considerable success, but in ways which differ from those unrealistic goals. Second, our basic understanding of
psychological development and how to foster it was inadequate to the task. This explains in considerable part the unrealism of attempting to overcome the deficit deriving from four years of experience in a summer or a year of compensatory education without altering the milieu. Third, the nature of the nursery-schooling available for deployment in the crash program of Headstart was poorly adapted for the compensatory effect called for by the goals.

In one sense, the evidence outlined here may be viewed as optimistic, perhaps too optimistic. It is one thing to say that most of the class and race differences now evident are not biologically inevitable, and it is quite another to say that reducing the deficits associated with poverty is easy. We lack basic knowledge of early intellectual and motivational development. Only recently have we begun to take seriously the hierarchical conception of such development and begun to describe the natural landmarks of achievement. An initial approximation of such landmarks exist now only for sensorimotor development (Uzgiris & Hunt, 1973) and for linguistic development during the preschool years (Brown, 1973). We know exceedingly little of what kinds of experience foster these successive landmarks and how each one is built upon earlier achievements. Even though we have a few instances where curricular suggestions have worked better than typical middle-class rearing, no suggestion we can make now is more than a hypothesis to be tested. Finding ways to get parents of poverty to utilize innovations in early education is another problem. From the experience of the Parent-and-Child Centers, it becomes clear that we have not even begun to evaluate programs in terms of the determinants of their success in eliciting parent cooperation and participation and to examine the factors responsible for success or failure. For many groups of parents any inferiority of their practices is hard to take, and resentment hampers cooperation with the program. Discovering ways of harnessing class, neighborhood, and ethnic pride to get parental cooperation in the improvement of early education along with discovering the kinds of experience which foster early intellectual and motivational development are the basic challenges for the behavioral and social sciences and professions concerned in early childhood education.
REFERENCES


Finch, F. H. 1946. Enrollment increases and changes in the mental level of the high school population. *Appl. Psychol. Monograph, 10*, 75.


