

The IQ controversy: A reply to Layzer

ARTHUR R. JENSEN

University of California, Berkeley

The publication of David Layzer's 'Science or superstition?' in the previous issue of this journal brings to 120 the total number of articles and books provoked by my article in the *Harvard education review* (Jensen, 1969a). A bibliography of these items appears in my recent book *Genetics and education* (Jensen, 1972c), the Preface to which also provides a chronicle of my involvement in 'the IQ controversy'. This book, together with my more recently published *Educability and group differences* (Jensen, 1973), actually give my detailed answers to practically all of the questions, criticisms, and issues raised in Layzer's article. I therefore urge readers who wish to gain a greater understanding of these matters, and of my own position concerning their educational implications, to delve into these books and see for themselves just what I am saying – completely and in context.

This is important because it is a common feature of so many of the criticisms of my position that they have had to misrepresent it and distort it, at times in ridiculous ways, in order to criticize it with the appearance of discrediting my main arguments. In this, Layzer's critique is no exception. Whatever the acumen that Layzer, as a physical scientist, might be able to bring to bear in this field if he were not so politically or ideologically involved, his article makes it all too obvious that he is very one-sided in the exercise of his critical judgment. In discussing the rationale and findings of studies which point to the strong involvement of genetic factors in the distribution of mental ability, Layzer assumes the posture of an extreme methodological blue-nose. Yet he shows total suspension of his critical powers when dealing with studies which he perceives (at times wrongly) as lending support to his 'environmentalist' and 'anti-hereditarian' attitudes.

I am not arguing with Layzer's political and social egalitarianism. But, I am saying that genetic equality of human abilities is an altogether untenable belief in view of the evidence we already possess, as untenable as the geocentric theory in astronomy or the doctrine of special creation in biology. I find nothing in Layzer's article that contradicts my main conclusions regarding the inheritance of mental

ability, and at times Layzer himself seems to acknowledge the importance of genetic factors (e.g., 'this result clearly indicates that genetic factors can play an important role in the development of cognitive skills'). Yet, for some reason, Layzer aims to denigrate or discredit evidence which attempts to achieve a greater precision in our knowledge of the relative influences of genetic and environmental factors (and their interaction) in the causation of human differences. The evidence we now possess leads me to the conclusion that in existing populations genetic factors significantly outweigh environmental influences in the distribution of human intelligence. The fact that the evidence is based on 'IQ tests' should not be construed to mean that such tests rank-order individuals much differently than if they were ordered in terms of parents', teachers', employers' or the 'man-in-the-street's' more subjective criteria for judging intelligence or 'brains'. If the concept of intelligence or the IQ were merely psychological esoterica, we can be fairly sure there would be no 'IQ controversy'.

Misrepresentations

Like so many of my critics, Layzer falsely attributes non sequiturs and absurdities to me and then boldly attacks them. Is it necessary to distort what I have actually said in order to find fault with it? For example, Layzer writes, 'such studies show, according to Jensen, that IQ differences are approximately 90 % genetic in origin'. I have never stated any figure as high as 90 % as an average. What in fact I have said in numerous articles is that in accounting for the causes of the differences among persons in IQ, genetic factors outweigh environmental influences by about 2 to 1. This is quite different from the 9 to 1 ratio implied by Layzer's figure of 90 %. If the broad heritability of IQ is about .80 (as my review of the evidence in 1969 led me to conclude), then the proportion of genetic to nongenetic (or environmental) variance is in the ratio of 4 to 1. Since the variance is derived from *squared differences*, the relative contributions of genetic and environmental effects to the actual *differences* in IQ would be in the ratio of $\sqrt{4}$ to $\sqrt{1}$, or 2 to 1.

Arguments as to whether the broad heritability (h^2) of IQ is .80 or some other value is another matter. Since by the definitional nature of h^2 there can be no one 'true' or constant value of h^2 for intelligence (or any other metric trait), it is fatuous to argue whether h^2 is .60 or .80 or .90 or some other value. But competent analyses of existing evidence find that most obtained values of broad h^2 fall in the range from .60 to .90 for various mental tests and various populations.

Layzer notes that I pointed out (Jensen, 1969b, p. 50) that the median correlations of .75 between the IQs of identical twins reared apart and .24 between IQs of unrelated children reared together are quite consistent with one another if the

heritability of intelligence is .75, since the correlation of .75 (an estimate of the genetic variance) plus the correlation of .24 (an estimate of environmental variance) totals .99, which comes very close to 1.00 or 100 % of the phenotypic variance. Actually, the theoretical discrepancy should be even more than .01, and probably closer to .10, since the correlation between unrelated children reared together does not include the within-families component of variance and the observed correlation is most likely inflated slightly by some degree of genetic correlation due to selective placement by the adoption agencies. In any case, sampling error for these correlations would exceed .01. Yet Layzer claims that I attribute this left-over .01 to genotype \times environment interaction! I have never drawn any such unwarranted and absurd conclusion. It is a figment of Layzer's, perhaps invented for the sake of criticism. Or perhaps he got it from Light and Smith (1969, p. 496), who also falsely attributed this absurdity to me and then proceeded to make it the keystone of an elaborate hypothetical analysis intended to demolish my incorrectly represented position. (The Light and Smith analysis, incidentally, proved fatally fallacious on other grounds as well; see Shockley, 1971a, 1971b; Light and Smith, 1971.)

Layzer complains that I ignore sex differences and color differences as sources of IQ variance, when in fact I have written a major research paper on sex differences and their interaction with racial differences (Jensen, 1971), and I have also explicated methods for studying the contribution of the social aspects of skin color to racial IQ differences (Jensen, 1969b, p. 241; 1970, pp. 150-151; 1973, pp. 222-227).

Layzer quotes the British geneticist Waddington to the effect that '... if one takes some particular phenotypic character such as body weight or milk yield, one of the first steps in an analysis of its genetic basis should be to try to break down the underlying physiological systems into a number of more or less independent factors'. Layzer comments: 'These views contrast sharply with those of Jensen ...' In reply, let me simply quote what I have written about the genetic analysis of mental abilities: 'A heritability study may be regarded as a Geiger counter with which one scans the territory in order to find the spot one can most profitably begin to dig for ore. Characteristics with low heritability are less likely to yield pay dirt. The reason, of course, is that all we have to work with, at least at the beginning of our investigation, is *variance*, and if what we are interested in is genetical analysis, we would like to know that some substantial proportion of the trait variance we are concerned with is attributable to genetic factors. So we should not belittle heritability studies; but they should be regarded as only the beginning rather than as the goal of our efforts in genetical analysis... a test score usually represents an amalgam of a number of psychological processes in each of which there are imperfectly correlated and genetically conditioned individual differences. Thus our

aim should be to fractionate our ability measurements so as to get at smaller and more unitary components of ability. This is the province of the differential psychologist, but it requires also the methods of experimental psychology. Factor analysis alone is not the answer' (Jensen, 1972b, p. 243). I go on to explain that factor analysis, as it has generally been used, has revealed common factors among tests which are already so complex as to permit nothing more than the fitting of a crude polygenic model, and I suggest greater experimental 'fractionation' of abilities in order to permit a more fine-grained and penetrating behavior-genetic analysis of them (Jensen, 1972b). Is this so very different from the views stated by Waddington? It is a fact that the heritability of a complex behavior or processes (like performance on an IQ test, or milk production in cows, or physical growth) can be determined without knowing anything about the underlying mechanisms. But as scientists, of course, we wish to attain an ever more complete understanding of the phenomenon. A substantial heritability index tells us that an understanding of individual differences in the trait in question must be sought in the organism's internal, genetically conditioned biochemical and physiological processes as well as in the external environmental influences acting throughout its development.

Misconceptions about heritability

Part of Layzer's criticisms of the heritability concept seems to be based on the disillusionment of his own assumptions as to what it means, or at least the meaning he thinks others give to it. But heritability was never intended to have some of these meanings, and I do not know any fully informed persons who have held the beliefs about heritability that Layzer so deplors. Heritability (h^2) is a technical term in genetics which refers to the proportion of the population variance in a phenotypic characteristic or measurement that is attributable to genetic variation. Narrow heritability includes only the additive part of the genetic variance, i.e., the part which 'breeds true' and largely accounts for the resemblance between parents and children. Psychologists are more interested in broad heritability, which includes all the genetic variation, the additive portion plus variance due to interaction between genes at the same loci on the chromosomes (called *dominance*) and interactions among genes at different loci (called *epistasis*). Broad heritability, which is what we are concerned with here, is sometimes referred to as the *coefficient of genetic determination* to distinguish it from narrow heritability. Broad heritability can take any value from 0 to 1. It is not a constant, but differs for different traits, different measurements, and in different populations. Its value can be estimated by a number of methods in quantitative genetics. Like any population statistic, it is subject to measurement error and sampling error. Since it is based

essentially on the analysis of variance, it can tell us nothing at all about the causes of the particular value assumed by the grand mean of the population. It only analyzes the variance (or squared deviations) *about* the grand mean. And it tells us what proportion of this total variance is *genetic* variance and what proportion is *nongenetic*, i.e., due to environmental factors of all kinds, to interaction and covariance of genetic and nongenetic factors, and to errors of measurement. Most estimates of the broad heritability of IQ in the European and North American populations on which we have reasonably good data fall in the range from .60 to .90, and most of these estimates are in the range from .70 to .80.

We could never determine heritability by studying a single individual, since heritability deals with *differences* among individuals. Each individual's own development from the moment of conception is the inextricable product of his genes and environment; both are as necessary as are length and width in determining the area of a rectangle. But it is possible to analyze the phenotypic variance of a population sample into additive components of variance attributable to the additive effects of genetic factors and environmental factors, to their nonadditive effects (called $G \times E$ interaction) and to the covariance (or correlation) between the genetic and environmental factors. The quantitative methods for doing this are explicated in textbooks on quantitative and population genetics. The most recent and sophisticated application of these techniques to psychological test data has been made by the British geneticists Jinks and Fulker (1970).

A major point in Layzer's argument is that IQ scores, in principle, are unsuitable for this type of analysis, since they do not constitute an absolute scale (which is distinguished by a true zero point and equal intervals) like height and weight. An IQ is actually a normalized standard score, which indicates an individual's standing, represented on a scale of deviates of the normal curve (multiplied by some arbitrary constant, such as 15 for the standard deviation of IQ), in relation to some reference population. In the case of the most widely used standardized tests, the reference population (or 'norms', as they are called) is a representative sample (usually nationwide) of individuals of the same age as the individual whose IQ we wish to determine. Thus there is no true zero point, and the IQ scale can be regarded as an interval scale only if we make the assumption that intelligence should have an approximately normal distribution in the population. Such measurements *can* meaningfully be subjected to the statistical techniques of analysis of variance, regression analysis, and correlation analysis – the principal tools of quantitative genetics. Since essentially all we are analyzing are the squared deviations from the mean, measurements on an absolute scale are quite unnecessary. (True, an absolute scale for intelligence would be a great advantage for some other purposes, such as studying the form of the growth curve of intelligence, but it is

not a necessary condition for quantitative-genetic analysis.) Absolute scale characteristics such as height and weight are often converted to deviation scores within age groups, thus making their scale properties essentially the same as the IQ scale, prior to a genetic analysis of these physical traits. Some such transformation of the original measurements is needed to take care of sex and age differences when the genetic analysis involves correlations between parents and children, siblings, cousins, etc. One might argue that deviation measurements are permissible when the original measurements on which they are based is an absolute scale, as in the case of height and weight. Granted, the measurements may be more reliable, or more unidimensional, but those are quite different and separate issues. The fact is that the methods of quantitative genetics work as well for deviation scales as for absolute scales. If this were not so, it is doubtful that some of the world's leading quantitative geneticists, such as Sewell Wright and John F. Jinks, would have undertaken genetical analyses of IQ scores. Textbooks on quantitative genetics do not limit the application of these methods only to characteristics that can be measured on an absolute scale. Many geneticists have written about the genetics of intelligence and found no fault with the scale properties of the IQ as regards its suitability for heritability analysis. Thus, it seems to me, a major pillar of Layzer's critique simply crumbles.

Experience with transformations of physical measurements to a scale equivalent to the IQ scale has shown me that when the sample size is reasonably large (i.e., 50 or more) and the measures are fairly continuous and unimodal, the difference between correlations (or results of analysis of variance) based on the original (absolute) measurements and on the deviation scores (equivalent to IQs) is practically nil. Now, if two or more studies are based on *different* tests in which the deviation scores are based on significantly *different* normative populations, you may average the components of variation (expressed as proportions of total variance) revealed by the genetical analysis, but what you cannot do is translate these proportions back into any scalar quantities. That is, the proportions of genetic and environmental variance can no longer be expressed in terms of number of IQ points difference (on average) between individuals attributable to genetic or to environmental factors (or to any other components). To do that, one would need to make the assumption that the tests and norms were the same in the two or more studies that had been averaged. Whether such an assumption is tenable is an empirical issue. But there is nothing wrong with averaging proportions of variance as a way of summarizing the central tendency of a number of studies, as long as it is realized that they are only proportions of variance and are not misinterpreted as scalar quantities.

Heritability and teachability

The fact that IQ has high heritability surely does *not* mean that individuals cannot learn much. Even if learning ability had 100 % heritability it would not mean that individuals cannot learn, and therefore the demonstration of learning or the improvement of performance, with or without specific instruction or intervention by a teacher, says absolutely nothing about heritability. But knowing that learning ability has high heritability does tell us this: If a number of individuals are all given equal opportunity – the same background, the same conditions, and the same amount of time – for learning something, they will still differ from one another in their rates of learning and consequently in the amount they learn per unit of time spent in learning. That is the meaning of heritability. It does not say that individuals cannot learn or improve with instruction and practice. It says that given equal conditions, individuals will differ from one another, not because of differences in the external conditions but because of differences in the internal environment which is conditioned by genetic factors. ‘Teachability’ presumably means the ability to learn under conditions of instruction by a teacher. If this is the case, then it is true that heritability has nothing to do with teachability. But was this ever really the question? Has anyone questioned the fact that *all* school-children are teachable? The important question has concerned *differences* in teachability – differences both among individuals and among subgroups of the population. And with reference to the question of *differences*, the concept of heritability is indeed a relevant and empirically answerable question.

We have heard it said that ‘teachability is not inversely related to heritability’. Such a statement simply ignores the central fact that heritability deals with differences. The degree to which equal conditions of teaching or instruction will diminish individual differences in achievement *is* inversely related to the heritability of the ‘teachability’ of the subject in question, and various school subjects probably differ considerably in heritability.

The fact that scholastic achievement shows lower heritability than IQ means that more of the variance in scholastic achievement is attributable to nongenetic factors than is the case for IQ. Consequently, we can hypothesize what the sources of the environmental variance in scholastic achievement are, and possibly we can manipulate them. For example, it might be hypothesized that one source of environmental variance in reading achievement is whether or not the child’s parents read to him between the ages of 3 and 4, and we can obviously test this hypothesis experimentally. Much of the psychological research on the environmental correlates of scholastic achievement has been of this nature. The proportion of variance indicated by $1-h^2$, if small, does in fact mean that the sources of environmental variance are skimpy under the conditions that prevailed in the population in which h^2 was estimated. It means that the *already existing* variations in environmental

(or instructional) conditions are not a potent source of phenotypic variance, so that making the best variations available to everyone will do relatively little to reduce individual differences. This is not to say that as yet undiscovered environmental manipulations or forms of intervention in the learning or developmental process cannot, in principle, markedly reduce individual differences in a trait which under ordinary conditions has very high heritability. By the same token, low heritability does not guarantee that most of the nongenetic sources of variance can be manipulated systematically. A multitude of uncontrollable, fortuitous microenvironmental events may constitute the largest source of phenotypic variance in some traits.

The heritability of individual differences and of group differences in scholastic performance in the total population are therefore relevant if we are at all interested in the causes of these differences. To say that heritability is trivial or irrelevant is to say also that the complement of heritability, $1-h^2$, or the proportion of variance attributable to nongenetic or environmental factors is also trivial. To dismiss the question of heritability is to dismiss concern with the causes of educational differences and their implications for educational practices. As I read it, what most educators, government officials, and writers in the popular press who discuss the present problems of education are in fact referring to is not primarily dissatisfaction with some absolute level of achievement, but rather with the large group differences in educational attainments that show up so conspicuously in our educational system – the achievement gaps between the affluent and the poor, the lower-class and the middle-class, one race and another, the majority and the minority, the urban and the suburban, and so on. Educational differences, not absolute level of performance, are the main cause of concern. Whether we like to admit it or not, the problem of achievement differences today is where the action is, where the billions of dollars of educational funds are being poured in, where the heat is on, and where the schools are being torn apart. Are we not trying to understand more about the causes of these differences?

It is mistaken to argue that heritability has no implications for the probable effects of environmental intervention. Since $1-h_c^2$ (h_c^2 is h^2 corrected for attenuation) is the proportion of trait variance attributable to nongenetic or environmental factors, the square root of this value times the *SD* of the 'true score' trait measurement gives the *SD* of the effect of existing environmental variations on the particular trait. For IQ this is about six points; that is to say, a shift of one *SD* in the sum total of whatever nongenetic influences contribute to environmental variance (i.e., $1-h_c^2$), will shift the IQ about six points. Thus, the magnitude of change in a trait effected by changing the allocation of the existing environmental sources of variance in that trait is logically related to its heritability. This applies, of course, only to existing sources of environmental variance in the population, which is all that can be estimated by $1-h_c^2$. It can

have no relevance to speculations about as yet nonexistent environmental influences or entirely new combinations of already existing environmental factors. With respect to IQ, I believe Carl Bereiter (1970) stated the situation quite correctly: 'What a high heritability ratio implies, therefore, is that changes within the existing range of environmental conditions can have substantial effects on the mean level of IQ in the population but they are unlikely to have much effects on the spread of individual differences in IQ within that population. If one is concerned with relative standing of individuals within the population, the prospects for doing anything about this through existing educational means are thus not good. Even with a massive redistribution of environmental conditions, one would expect to find the lower quarter of the IQ distribution to be about as far removed from the upper quarter as before' (p. 288).

Genotype \times environment interaction

Layzer makes much of the possibility of interaction of genetic and environmental factors. Interaction in this case means that the population variance of the phenotypic measurements is composed in whole or in some part of the nonadditive effects of genetic and environmental factors. The existing models of heritability analysis take such interaction into account and are capable of estimating the proportion of variance attributable to such interaction. With respect to IQ, the fact is that this interaction component is either nonexistent or so insignificant as to be undetectable in the existing data. If it were of substantial magnitude, it would easily show up with the present methods of analysis, which are quite capable of detecting other forms of interaction, such as dominance. In reading Layzer, one might easily get the impression that there is a lot of $G \times E$ interaction but that our models are unsuited to detecting it. Not so (see Jinks and Fulker, 1970). The fact that the genetic model for heritability is an *additive* model (as is all analysis of variance) does not mean that all of the components are forced into being either G or E ; some of the components can be nonadditive functions of G and E . True, geneticists usually try to account for as much of the total variance as possible in terms of the strictly additive effects of G and E and will often make some scale transformation of their original measurements in order to minimize or eliminate the nonadditive components of variance. But this is unnecessary for IQ scores, which show little or no $G \times E$ interaction; the additive model fits IQ data about as well as it fits data on physical characteristics like height and weight. One of the impressive facts about genetical analyses of the IQ is how much it behaves like measures of continuous physical traits.

Experimentally, psychologists have not discovered any teaching methods of environmental manipulations which interact with IQ in such a way as to wipe out differences in learning between individuals differing in IQ. The search for aptitude \times

training interactions, or ATI for short, has become a popular area for research in educational psychology. What ATI means, simply, is that no single instructional method is best for everyone, and that optimal performance will result only by matching a diversity of instructional methods with the diversity of individual's aptitudes. If Bill and John are both taught by method *A* and Bill does much better than John, perhaps there is a different teaching method, *B*, that will permit John to learn as fast as Bill. That is the hope of ATI researchers. The only trouble so far has been that when you find a method *B* which boosts John's performance a little, it usually does so even more for Bill. Bracht (1970) recently reviewed a large number of studies in the ATI field which met certain methodological and statistical criteria to permit rigorous evaluation, and he found that out of 90 studies that were specifically designed to yield aptitude \times treatment interactions of the kind that would solve the performance difference between Bill and John, only five yielded such an $A \times T$ interaction, and none of these aptitude differences was of the IQ variety – they were personological variables unrelated to intelligence. Bracht says a number of interesting and important things:

'When a variety of treatment stimuli, especially conditions not controlled by the experimenter, are able to influence performance on the dependent variable, it is unlikely that a personological variable can be found to produce a disordinal interaction with the alternative treatments . . . Success on a combination of heterogeneous treatment tasks is predicted best by measures of general ability [i.e., IQ tests], and the degree of prediction is about equally high for alternative treatments' (p. 636). 'The degree of task complexity may be a major factor in the occurrence of ATI. Although the treatment tasks for most of the 90 studies were classified as controlled, the treatments were generally relatively complex tasks. Conversely, four of the five experiments with disordinal interactions [ATI] were more similar to the basic learning tasks of the research laboratory' (p. 637). 'Despite the large number of comparative experiments with intelligence as a personological variable, no evidence was found to suggest that the IQ score and similar measures of general ability are useful variables for differentiating alternative treatments for subjects in a homogeneous age group. These measures correlate substantially with achievement in most school-related tasks and hence are not likely to correlate differentially with performance in alternative treatments of complex achievement-oriented tasks' (p. 638).

Such findings give little cause for optimism in finding new teaching methods that will overcome the large scholastic achievement differences that are so evident in our schools and are highly related to IQ.

Heritability and group differences

I have often been falsely accused of claiming that the high heritability of IQ inevitably means that the mean differences in IQ between social class groups and racial groups must be due to genetic factors. I have never made this incorrect inference. What I have said is this: While it is true, indeed axiomatic, that heritability *within* groups cannot establish heritability *between* group means, high *within* group heritability increases the *a priori* likelihood that the *between* groups heritability is greater than zero. In nature, characteristics that vary genetically *among* individuals within a population also generally vary genetically *between* different breeding populations of the same species. Among the genetically conditioned traits known to vary between major racial groups are body size and proportions, cranial size and cephalic index, pigmentation of the hair, skin, and eyes, hair form and distribution on the body, number of vertebrae, fingerprints, bone density, basic metabolic rate, sweating, fissural patterns on the chewing surface of the teeth, numerous blood groups, various chronic diseases, frequency of dizygotic (but not monozygotic) twinning, male/female birth ratio, ability to taste phenylthiocarbamide, length of gestation period, and degree of physical maturity at birth (as indicated by degree of ossification of cartilage). In light of all these differences, Spuhler and Lindzey (1967) have remarked '... it seems to us surprising that one would accept present findings in regard to the existence of genetic anatomical, physiological, and epidemiological differences between the races ... and still expect to find *no* meaningful differences in behavior between races' (p. 413). The high within-groups heritability of certain behavioral traits, such as intelligence, adds weight to this statement by Spuhler and Lindzey.

In fact, it is quite erroneous to say there is no relationship whatsoever between heritability *within* groups and heritability *between* group means. Jay Lush, a pioneer in quantitative genetics, has shown the formal relationship between these two heritabilities (1968, p. 312), and it has been recently introduced into the discussion of racial differences by another geneticist, John C. DeFries (1972). This formulation of the relationship between heritability *between* group means (h_B^2) and heritability *within* groups (h_W^2) is as follows:

$$h_B^2 \simeq h_W^2 \frac{(1-r) \rho}{(1-\rho) r}$$

- where:
- h_B^2 is the heritability between group means.
 - h_W^2 is the average heritability *within* groups.
 - r is the intraclass correlation among *phenotypes* within groups (or the square of the point biserial correlation between the quantized racial dichotomy and the trait measurement).
 - ρ is the intraclass correlation among *genotypes* within groups, i.e., the within-group genetic correlation for the trait in question.

Since we do not know ρ , the formula is not presently of practical use in determining the heritability of mean group differences. But it does show that if, for a given trait, the genetic correlation among persons within groups is greater than zero, the between-group heritability is a monotonically increasing function of within-groups heritability. And the probability that a phenotypic mean difference between two groups is in the same direction as the genotypic mean difference is greater than the probability that the phenotypic and genotypic differences are in opposite directions. I know no principles of genetics that would rule out the possibility of determining the heritability of differences between group means. If this question is unresolvable in the sense, say, that perpetual motion is impossible, Layzer should be able to spell out the laws of nature that make it so. To say it is possible in principle, however, is not to say it is easy in practice. The methods would have to differ from those used for determining the heritability of individual differences, just as the method for determining the temperature of a distant star must differ from that for measuring the temperature in the kitchen stove. The science of astronomy would never have advanced beyond star gazing if astronomers had applied as little imagination and ingenuity to solving problems in their field as Layzer seems to insist be applied in genetics.

The storm of criticisms and ideological protests directed at me has been a result of my expressing serious doubts that the observed mean IQ difference between whites and blacks in the U.S., and between social class groups, is entirely explainable in terms of culture-bias in tests, unequal educational opportunities, social discrimination, and other environmental influences. My position is that there is now sufficient evidence seriously to question the 100 percent environmental theories of the white-black intelligence difference. Are there any responsible scientists today who claim that this position can be ruled out on the basis of evidence or ruled out *a priori* by any principle of genetics? How many scientists today express little or no doubt that all of the racial IQ difference is attributable to environment? And on what evidence do those who claim no doubt base their certainty? I have not found any 100 % environmental theory which can explain the facts or which stands up when its major premises are critically examined in the light of evidence. Therefore, I regard this issue scientifically as an open question which can be eventually answered in a scientific sense only if we are willing to consider all reasonable hypotheses. It is a reasonable and potentially testable hypothesis that genetic factors are involved in the average white-black IQ difference. My study of the research evidence bearing on this question at present leads me to believe that a preponderance of the evidence is more consistent with a genetic hypothesis, which, of course, does not exclude the influence of environment (Jensen, 1973).

Means versus medians

Layzer complains about my use of the median (i.e., the middle value in a distribution) for summarizing the correlations obtained in numerous studies of various kinship groups (Jensen, 1969a, Table 2, p. 49). I had noted that these median correlations between IQs for various degrees of kinship come very close to the values one should expect from a polygenic model of inheritance, and they are the basis for the conclusion that genetic factors predominate as a cause of IQ differences in the populations in which these kinship correlations were obtained. Layzer believes that the median is not the proper statistic for indicating the central tendency of a number of correlations obtained in various studies. He suggests that instead of the median one should use the mean of the various obtained correlations – a weighted mean, with each correlation entering into it weighted inversely to its probable error. This is a correct and standard way for combining statistics, and in general I agree with it. Weighting the correlations by their standard error, *SE*, (or the probable error, which is $.67 SE$) surely makes sense, since the *SE* indicates the precision or reliability of the sample estimates of the population value, and of course we would want to give more weight to the more reliable values. On the other hand, an argument can be made in this particular case for using the median instead of a weighted mean. The median, of course, is least affected by extreme or atypical values. Since the kinship correlations reported in the literature are based on a variety of tests, some of which are scholastic achievement tests or tests of very narrow and special abilities rather than tests of general intelligence, a decision has to be made concerning which tests to include in the collection of correlations of which we wish to represent the central tendency. Errors of judgment on this point would little affect the median but could markedly affect the weighted mean, particularly if the correlation for an atypical test or population were based on a very large size sample. As an example, one of the largest sets of twin data ever collected consists of a nationwide sample of monozygotic (MZ) and dizygotic (DZ) twins who as high school seniors entered the National Merit Scholarship competition and took the set of scholastic achievement tests which serve as part of the basis, along with high school grades, for picking the winners. In the first place, tests of scholastic achievement generally show much higher correlations between any children reared together (even when they are genetically unrelated) than do intelligence tests, and the difference between MZ and DZ twin correlations is much less for scholastic tests than for IQ tests. This difference in the case of the National Merit Scholarship data is further diminished by the fact that poor students do not enter the competition, and since DZ twins are less likely to be alike than MZ twins, there will be more instances where only one member of a DZ twin pair will get into the National Merit Scholarship screening process than in the case

of MZ twins. This differential selection bias in the MZ and DZ twin samples makes the twin correlations (and consequently any estimates of heritability derived from them) atypical. The probable error, however, is smaller than for any other study, so a weighted mean including the National Merit Scholarship correlations could be quite distorted. The median would be much less distorted.

Also when it comes to weighting the various entries, one could make a case for weighting correlations in terms of more than just their probable error. Why not weight the correlations or heritabilities derived from any given study in terms of the degree to which the particular test used in the study is loaded with *g*, that is, the extent to which the test, when factor analyzed with other tests of intelligence, shares common variance with the other tests? Tests that have less in common with what we are calling intelligence (*viz.*, the *g* or general factor common to all complex mental tests) would be given less weight in the composite weighted mean. Why not assign weights according to the representativeness of the sample? Should a heritability estimate based on college graduates be given as much weight as one based on a representative sample of elementary school children? Should we weight in terms of the degree to which the various sample means and variances approach the population values for the tests used in the various studies? How about differentially weighting studies that differ in the degree to which they meet certain assumptions that underlie the methods for estimating heritability, such as equality of the total variances in both the MZ and DZ twin samples? We can see that there can be many other criteria for weighting besides just the probable error of the obtained correlations. I do not advocate such elaborate weighting, because I believe it can introduce too much subjectivity and, since many of the weights themselves are subject to error, would tend to lower our confidence in the composite. (Weighting by the *SE* alone, however, does not have this drawback.) All things considered, therefore, I feel that with these data there is apt to be less risk of distortion in the median than for any other measure.

Anyway, it should be interesting to see how much difference it would make if we used weighted means instead of medians. Layzer's readers may have been led to believe that the weighted means would give a quite different picture from that provided by the medians. I have obtained both medians and weighted means of all the reported kinship correlations that I can find in my reprint files which are based on some kind of general intelligence test.¹ (I have excluded purely scholastic achievement tests.) The individual correlations were weighted by their standard errors. (In accord with the standard statistical procedure for averaging correlations coefficients,

1. It would take up too much space to list here all of the published sources of these kinship correlations. The writer will provide

the list of references, keyed to Table 1, to anyone requesting it.

Table 1. Comparison of median correlation and weighted mean correlation from various kinship studies

Degree of relatedness	Number (N) of correlations	Number of subject pairs	Median correlation	Weighted* mean correlation	SD of N correlations	SE of mean correlation	Correlations for physical measures**	
							Height	Weight
Monozygotic twins								
Reared together	19	1626	.87	.87	.088	.024	.96	.93
Reared apart	9	122	.74	.76	.101	.091	.94	.88
Dyzygotic twins reared together	19	1384	.55	.55	.075	.026	.47	.59
Full siblings reared together	7	1935	.53	.55	.080	.022	.50	.57
Parents and children reared by their natural parents	6	3282	.47	.47	.076	.017		
Adopting parents and adopted children	6	1181	.21	.23	.100	.029		
Unrelated children reared together	9	385	.25	.32	.174	.051	-.07	.24

* The separate correlation coefficients, r , entering into each mean were transformed by Fisher's Z transformation and were then weighted inversely by the standard error of the transformed correlation, i.e., $1/N-3$; the weighted mean Z then was transformed back to the correlation coefficient, r .

** From Burt (1966).

r , these were first transformed to Fisher's Z , then weighted by the inverse of the SE , then averaged, and finally, transformed from Z back to r .) Table 1 shows the results. We see that the medians and weighted means are quite similar (in fact, they correlate .995). Also shown is the standard deviation (SD) of the correlations from the various studies. If all studies represented samples from the same population (of persons and tests), we should expect the SD of the obtained correlations to be close to the standard error of the mean correlation. The fact that the SD is slightly larger than the SE indicates that the correlations obtained in the various studies are more variable than we should expect if all had used the same test and had sampled from the same population.²

The last two columns in Table 1 show the correlations for physical measures. It can be seen that they follow the same pattern as the correlations for the mental tests. The fit of the data to the values expected according to a simple polygenic model is remarkably close. It suggests that the heritabilities of measures of intelligence and of body weight are very similar and consistent with a broad heritability of between .70 and .80. I do not know of any strictly environmental theories that can explain this pattern of correlations as well as does the polygenic model of inheritance. These correlations substantiate the conclusion of a greater genetic than environmental determination of individual differences in IQ. No geneticist who has studied such evidence has, to my knowledge, drawn the opposite conclusion, and I have searched all the up-to-date textbooks of genetics that deal with this subject. In regard to this evidence, Layzer himself writes, 'these findings show that IQ is strongly influenced by both genetic and environmental factors'. Though the genetic factors are in fact predominant, even if they were not, they eventually would become so as we achieved more and more equality of environmental conditions, cultural and educational opportunities, and the like, thereby reducing the environmental variance.

Twin differences and environmental differences

Layzer points out that the IQ differences between MZ twins reared apart show a fairly substantial correlation with ratings of the amount of difference in the environments in which they grew up. This point seems to be made with the idea that it somehow contradicts the high heritability of IQ as indicated by the high correlation

2. Since the SE of the weighted mean correlations shown in Table 1 is the SE for the total number of kinship pairs, the SE for any one of the studies entering into the weighted mean would of course be considerably larger. A rough approximation to the average SE

for single studies would be given by the value of SE in Table 1 multiplied by $1/N$, where N is the number of correlations. These values of $SE/1/N$, interestingly, differ but little from the SD of the N correlations.

between MZ twins reared apart. Since both members of a pair of MZ twins possess exactly the same complement of genes, any difference between them must of necessity be due to nongenetic causes. Thus it should not be at all surprising that the magnitude of the difference between their IQs is correlated with differences in the environmental conditions to which they were exposed. This fact in no way alters the fact that the nongenetic variance is quite small (about .20 to .30 of the total IQ variance). Moreover, by no means all of this nongenetic variance is attributable to what we ordinarily think of as 'environmental.' I have written in detail on this point (Jensen, 1972). Had Layzer carefully read my article, I doubt that he would have used the correlation between MZ twins' IQ differences and their environmental differences as part of his argument because this evidence actually weakens the case for the importance of social-psychological factors as a cause of IQ differences.

Information processing versus IQ

Layzer would prefer a measure of information processing capacity instead of the traditional IQ. Quantity of information can be measured on an absolute scale in terms of *bits*. (A *bit* [for binary digit] is the amount of information necessary to resolve two equally probable alternatives; it is equivalent to the minimum number of binary questions [answerable with Yes or No] needed to reduce uncertainty to zero. The number of *bits* is the logarithm, to the base 2, of the number of alternatives.) Such absolute measures have certain decided advantages in scientific research. Psychologists surely would welcome an instrument that measured a person's information processing on an absolute scale. Intelligence tests that involve problem solving and judgment are most likely measures of information processing capacity. The only trouble is that the items or problems that comprise such tests are so complex that we have no way, at least at present, of directly quantifying their informational content. The item difficulty of, say, Raven's Progressive Matrices (a nonverbal reasoning test) is probably highly correlated with the number of *bits* of information contained in the items. If we could determine the *bits* for every Raven item, it would be a boon to research in differential and developmental psychology. But would it change any of our main conclusions about the heritability of individual differences in *g* (which the Raven test largely measures)? I doubt it. I believe that information processing capacity is the essence of *g*, the general intelligence factor.

Layzer points to Piaget's conceptions of mental development and intelligence as being consistent with his idea of information processing. I agree. But let it be noted that Piaget has devised various special tests with which to study this information processing capacity, and when these Piagetian tests are given to large samples of children and are factor analyzed along with conventional tests of intelligence (e.g.,

the Stanford-Binet, the Wechsler tests, Raven's matrices, Kohs block designs, etc.), the Piagetian tests show high correlations with the other tests and are most highly loaded on the *g* factor; they tap little if any other source of variance not found in the conventional tests (Vernon, 1956). Moreover, the Piagetian tests show about the same magnitude of average differences between social classes and racial groups in California school children as are found with conventional IQ tests (Tuddenham, 1970).

When laboratory techniques have been specially devised to permit the actual measurement of information processing capacity in terms of time per *bits*, as in highly precise measures of visual information processing and of choice reaction time to differing amounts of information, quite striking social class and Negro-white differences have been found in the expected direction (Bosco, 1970; Noble, 1969; studies reviewed by Jensen, 1973, pp. 322-329).

The current denigration of the standard intelligence tests is a part of the attempt to minimize the significance of the evidence for a substantial genetic component in the variance on such tests; the scores on these tests are known to be correlated with educationally, occupationally, and socially significant criteria to about the same degree in different racial groups in the U.S.

Contrary to the popular mythology in this field, it is very difficult to find any objective evidence of culture bias that could account for social class and racial differences in performance on current standard tests of intelligence, even those, like the Peabody Picture Vocabulary Test (PPVT), which give the appearance of being highly culture-loaded. They may be culture-loaded, but there is no evidence we have been able to find that the culture-loading is what differentially affects the performance of Negro and white children. Difference in mean score cannot be a criterion of culture bias. One must seek other evidence. We have examined several types of evidence of culture-bias in the PPVT and Raven's Progressive Matrices. These studies have involved very large samples of Negro and white children in several California school districts.

We find that the rank order of the percent, *p*, passing each item is virtually the same for Negroes and whites. The *correlations* between the *p* values of Negroes and whites on these tests are all above .95, averaging .97. In this respect, the two racial groups are more alike than are boys and girls within each race. In other words, the cultural biases in the test are more apparent with respect to *sex* differences than with respect to *race* differences. (The sexes do not differ appreciably in mean score, however, while the racial groups differ about one standard deviation, or 15 IQ points, on the average.)

The matrix of item intercorrelations and the factor structure of these tests is not significantly different for white and Negro samples when these are roughly matched

for mental age or total score. These properties of the data, for example, do not in the least distinguish between 4th grade white children and 6th grade Negro children. Yet they distinguish between 5th grade and 6th grade Negro children, and between 5th grade and 6th grade white children. A culture-bias hypothesis would predict greater Negro-white differences than adjacent grade differences in item intercorrelations. The findings, on the other hand, are more consistent with a developmental lag hypothesis.

In multiple-choice tests, such as the PPVT and Raven, there is no systematic or significant racial difference in the choice of distractors on those items that are answered 'wrong'. A special scoring key was made up so as to score as correct whatever response is given by the largest number of children in the Negro sample. When the tests are scored by this key, the Negro sample still averages lower than the white sample.

Scales based on subgroups of items which discriminate either *least* between Negroes and whites or discriminate *most* are correlated with each other over .90 (approximately the reliability of the test), showing that the two types of items are measuring the same ability.

The intelligence tests show essentially the same size of correlation with scholastic achievement in Negro and white samples. When scholastic achievement is 'predicted' by a multiple regression equation comprised of several intelligence tests, adding race (white vs. Negro) to the multiple prediction equation does not increase the multiple correlation with scholastic achievement. The predictive validity of the IQ test is the same for Negroes and whites. Negroes and whites with the same IQ perform equally well in school.

In short, none of our analyses reveals any racial differences other than the number of items gotten right. There seems to be no good reason to believe that these tests behave any differently for Negroes than for whites.

The sibling correlations on 16 ability tests were examined in large Negro and white samples. They are very similar, as indicated by a correlation of .71 between the sibling correlations on each test for Negroes and whites. The average difference between siblings on each test does not differ significantly for Negroes and whites.

When estimates of the heritability (i.e., the proportion of genetic variance in test scores) of the various tests are correlated with the magnitude of the mean white-Negro difference on the tests, the correlation is positive (.80 for whites, .61 for Negroes). In other words, those tests which are *least* sensitive to environmental influences (i.e., high heritability) in general show the *largest* white-Negro differences, and those tests which are *most* sensitive to environmental influences (i.e., low heritability) show the *smallest* Negro-white differences. This outcome is just the opposite of what one would expect from a culture-bias or environmental hypothesis of the

cause of the racial difference. This study has been repeated by other investigators using a different set of tests, and the results are essentially the same, i.e., a strong positive correlation between tests' heritability and the magnitude of the white-Negro difference (for details see Jensen, 1973).

Those who claim culture bias in current widely used tests, it seems to me, are obligated to produce some objective evidence that such bias in fact exists. I have found no evidence that it does, at least in the well-known tests we have studied.

Misinterpretation of Skodak and Skeels

Layzer cites the famous study by Skodak and Skeels (1949) as if it contradicted my position regarding the heritability of IQ. For readers who might be misled into believing that the findings of this study are inconsistent with a genetic theory of intelligence and with the evidence on heritability, a brief review of it is in order.

Layzer's use of the Skodak and Skeels study is typical. The study is often held up by 'environmentalists' as an example of evidence which supposedly contradicts the high heritability of intelligence. The fact that the adopted children turned out to have considerably higher IQs than their biological mothers is thought to constitute a disproof of the conclusion from many heritability studies that genetic factors are more important than environmental factors (in the ratio of about 2 to 1) in the causation of individual differences in IQ. If about 80 percent of the IQ variance is attributable to genetic factors, the 20 percent of the variance due to environmental differences can be thought of as a more or less normal distribution of all the effects of environment on IQ, including prenatal and postnatal influences. This distribution of environmental effects would have a standard deviation of about 7 IQ points, since the total variance of IQ in the population is $15^2 = 225$ and the 20 percent of this which is attributable to environment is $.20 (225) = 45$, the square root of which gives $SD = 6.71$. Is there anything in the Skodak and Skeels data that would contradict this conclusion? Skodak and Skeels based their study on 100 children born to mothers with rather low IQs (a range from 53 to 128, with a mean of 85.7, SD of 15.8). The children were adopted into what Skodak and Skeels described as exceptionally good, upper-middle class families selected by the adoption agency for their superior qualities. Of the 100 true mothers, 63 were given the 1916 form of the Stanford-Binet IQ test at the time of the adoption. Their children, who had been reared in adoptive homes, were given the same test as adolescents. The correlation between the mothers' and children's IQs was .38. Layzer notes that the IQs of the adopted children average about 20 points higher than the IQs of their true mothers. However, the difference between the mothers' and children's IQs is not really the relevant question. It is on this point that the interpretation of this study has so often

been misleading. What we really want to know is, how much do the children differ from the IQs we would predict from a genetic model?³ Using a simple model (provided by Crow, 1971, p. 157), which assumes that the children represent a random selection of the offspring of mothers with a mean IQ of 85.7 and that the children are reared in a random sample of homes in the general population, the children's average predicted IQ should be 95. In fact, however, their average IQ turns out to be 106, or 11 points higher than the predicted IQ. If 20 percent of the IQ variance is environmental, and if one standard deviation of environmental influence is equivalent to about 7 IQ points, then it might be said the Skodak and Skeels children were reared in environments which averaged 11/7ths or about 1.6 standard deviations above the average environment of randomly selected families in the population. This would be about what one should expect if the adoption agency placed children only in homes they judged to be about one standard deviation or more above the average of the general population in the desirability of the environment they could provide. From what Skodak and Skeels say in their description of the adoptive families, they were at least one standard deviation above the general average in socioeconomic status and were probably even higher in other qualities deemed desirable in adoptive parents. So an eleven-point IQ gain over the average environment falls well within what we should expect, even if environmental factors contribute only 20 % of the IQ variance. In other words, this 11 points is well within the reaction range of phenotypic IQ, given a broad heritability of .80. But this 11 IQ points of apparent gain is more likely to be an overestimate to some extent, since these children, it should be remembered, were selected by the agency as suitable for adoption. They were not a random selection of children born to low IQ mothers. Many such children are never put out for adoption. (Most of the children were illegitimate, and as indicated in Leahy's, 1935, study, illegitimate children who become adopted have a higher average IQ than illegitimate children in general or than legitimate children placed for adoption.) Even so, it is interesting that Skodak and Skeels found that the 11 adopted children whose true mothers had IQs below 70 averaged 25 points lower than the 8 adopted children whose true mothers had IQs above 105. There are also certain technical, methodological deficiencies of the

3. This genetic prediction is sometimes made incorrectly by basing it on all 100 children, while actually we can make a prediction only for the 63 children whose true mothers' IQs were known. The model assumes (a) test reliability of .90; (b) an 'age attenuation' of .95 (due to the fact that the mothers and children are widely separated in age and the correlation between the IQs of the same per-

sons tested that many years apart is .95 after correction for immediate test-retest unreliability); (c) narrow heritability of .71 (the estimate of Jinks and Fulker, 1970, p. 342); the narrow heritability is used when predicting offsprings' values from parents' values; (d) random mating (since the mothers were unmarried and nothing is known about the IQs of the true fathers).

Skodak and Skeels study which make the basic data questionable; these deficiencies were trenchantly pointed out many years ago in critiques by Terman (1940, pp. 462-467) and McNemar (1940). But the Skodak and Skeels study, such as it is, can be seen to be not at all inconsistent with a heritability of .80 for intelligence.

To assume that the same 11-point IQ gain over the predicted value would have occurred if the biological mothers had been Negro instead of being white (but with exactly the same IQs) would be an unwarranted inference. It is unwarranted because according to the genetic model or prediction equation the Negro children would regress toward the Negro population mean IQ of about 85, rather than toward the white mean IQ of 100. Thus the predicted IQ gain of the adopted Negro children under the same environmental conditions would be some 10 IQ points less than the 11 IQ points gain for white children. If the Negro-white population difference in IQ is largely genetic, then a genetic model with dominance will predict regression of individual IQs to different population means for Negro and white children. The truth or falsity of this prediction is what we would like to know. The study of cross-racial adoptions might help to elucidate the matter. Since cross-racial adoptions are not hard to find, it is interesting that the environmentalists who go on citing the rather weak Skodak and Skeels study have never investigated similar data based on Negro children. It would be considerably more relevant.

A one-sided critical stance

As I noted earlier, Layzer suspends his critical judgment when citing those studies which he apparently believes support his position. There are many examples of this in his article, but at least three instances merit some comment, since they involve relatively recent publications which readers may not have had the chance to evaluate for themselves.

Layzer, being as methodologically puritan as possible in judging the evidence for the heritability of IQ, states '... measurements unaccompanied by error estimates have no scientific value'. If he had determined the error estimates of the data so which he was referring, he would have found the inferences based thereon to be highly significant (e.g., Jensen, 1967; 1972, pp. 294-306; Jinks and Fulker, 1970). On the other hand, all the points Layzer refers to in Scarr-Salapatek's study were presented by Scarr-Salapatek (1971) without any error estimates or tests of statistical significance. Furthermore, when the proper error estimates are made, it turns out that all the 'evidence' in Scarr-Salapatek's study regarding the comparative heritabilities of intelligence test scores in white and Negro samples, and in lower- and middle-class groups, is completely lacking in significance. The study has been subjected to a detailed examination by two leading quantitative geneticists (Eaves and

Jinks, 1972). Here is what they conclude about this study: 'On purely theoretical grounds, therefore, we suggest that this particular experimental design, with the small samples available, could not be expected to lead to the conclusions which were drawn and indeed could only be drawn from it by omitting proper tests of significance.' So large are the standard errors in Scarr-Salapatek's study that, as Eaves and Jinks point out, '... the data cannot even support the well-established conclusion that there is a genetical component of individual differences in intelligence'. Also, '... there is no evidence that the size of any heritable component depends on race or social advantage. This finding contradicts the main conclusion of Dr. Scarr-Salapatek's analysis which is based on a comparison of the numerical values of the correlations'. Finally, 'there is certainly no evidence in Scarr-Salapatek's studies that the proportion of genetical variation in either verbal or nonverbal IQ depends on race or social class'.

Layzer refers to Heber's Milwaukee Project as an example of the extreme plasticity of intelligence. He quotes Scarr-Salapatek's description of it, to the effect that IQs of ghetto children born to especially low-IQ mothers showed an enormous IQ gain of 37 points over a control group as a result of environmental intervention. It is unlikely that Layzer has critically examined this study, for there have been no published reports of it except for stories in the popular press, and the authors of the study have apparently not been willing to make technical reports of it available to other researchers in this field who have requested it, myself included.⁴ However, Professor Ellis B. Page, an expert in psychometrics and research methodology, managed to secure a detailed report of this study from the agency which funded it (and which no longer makes the report available). Page has subjected this report of the Milwaukee Study to detailed scrutiny (1972), and his findings should be of interest to anyone who, with Layzer, claims to insist upon methodological purity. Page's critique certainly leaves one with a markedly different impression of the study, and with a much greater skepticism, than is prompted by the sensationally optimistic reports appearing in the popular press. Page concludes: 'The Milwaukee Project, then, is here viewed as deficient on three counts: biased selection of treatment groups, contamination of criterion tests; and failure to specify the treatments. Any one of these would largely invalidate a study. Together, they destroy it.'

Layzer uncritically refers to an 'incisive critique of Jensen's [*HER*] article' by Deutsch (1969). In this 'critique' Deutsch claimed that some 17 errors were turned up in a casual perusal of my article (p. 524), and elsewhere he claimed that my article contained '... fifty-three major errors or misinterpretations, all of them uni-

4. Since this article went to press, I have, received a copy of the report from Dr. Heber's

office, some nine months after requesting it.

dimensional and all of them anti-black'. This claim is baseless and defamatory. It took 22 months of repeated prodding by the American Psychological Association's Committee on Scientific and Professional Ethics and Conduct to extract an itemized list of these 53 purported 'errors' from Deutsch. In view of all the efforts by ideological environmentalists to discredit my *HER* article, one wonders why Deutsch's list of 53 'major errors' has not gotten beyond the Ethics Committee and found its way into print. Considering the extreme pressure Deutsch was under from the Ethics Committee either to make a retraction of his defamatory claim or to produce a list of the '53 errors', it is most instructive, and I might add most flattering to my *HER* article, to see what Deutsch's list of '53 errors' actually consists of. It utterly fails to support his claim. I wish it were published, but since it is not, I will gladly send it to all who request it. Readers can judge for themselves the quality of Deutsch's 'incisive critique', to use Layzer's words.

One could go on noting other deficiencies in Layzer's critique, but many of his points are long since discredited arguments that would be apparent to most readers familiar with this literature; most of the issues are treated in more general terms in my other writings (see References). As to Layzer's ideological-political brand of environmentalism, I will make no comment here. My own position concerning the broader educational, societal, and ethical aspects of the genetical study of human differences has been amply expressed in numerous other articles (for a complete bibliography, see Jensen, 1972c, pp. 365-369).

The overwhelming fact is that the scientific world no longer presents a consensus of environmentalism to the public,⁵ and articles such as Layzer's will do nothing to restore the appearance of consensus which Layzer and his likes are so disturbed to see undone.

5. See the resolution on behavior and gene-

tics in the *Amer. Psychol.*, 1972, 27, 660-661.

REFERENCES

- Bereiter, C. (1970) Genetics and educability: Educational implications of the Jensen debate. In J. Hellmuth (Ed.), *Disadvantaged child*. Vol. 3. *Compensatory education: A national debate*. New York, Brunner-Mazel. Pp. 279-299.
- Bosco, J. J. (1970) Social class and the processing of visual information. Final Report, Project No. 9-E-041, Contract No. OEG-5-9-325041-0034(010). Office of Education, U.S. Dept. of Health, Education and Welfare.
- Bracht, G. H. (1970) Experimental factors related to aptitude-treatment interactions. *Rev. educ. Res.*, 40, 627-645.
- Burt, C. (1966) The genetic determination of differences in intelligence: A study of monozygotic twins reared together and apart. *Brit. J. Psychol.*, 57, 137-53.
- Crow, J. F. (1970) Do genetic factors contribute to poverty? In Allen, V. L. (Ed.), *Psychological factors in poverty*. Chicago, Markham. Pp. 147-160.
- DeFries, J. C. (1972) Quantitative aspects of genetics and environment in the determination of behavior. In Lee Ehrman, G. S. Omenn, and E. W. Caspari (Eds.), *Genetics, environment and behavior: Implications for educational policy*. New York, Academic Press. Pp. 5-16.
- Deutsch, M. (1969) Happenings on the way back to the forum: Social science, IQ, and race differences revisited. *Harv. educ. Rev.*, 39, 523-557.
- Eaves, L. J., and Jinks, J. L. (1972) Insignificance for differences in heritability of IQ between races and social classes. *Nature*, 240, 84-87.
- Jensen, A. R. (1967) Estimation of the limits of heritability of traits by comparison of monzygotic and dizygotic twins. *Proceedings of National Academy of Science*, 58, 149-157.
- Jensen, A. R. (1969a) How much can we boost IQ and scholastic achievement? *Harv. educ. Rev.*, 39, 1-123.
- Jensen, A. R. (1969b) Reducing the heredity-environment uncertainty, *Harv. educ. Rev.*, 39, 449-483.
- Jensen, A. R. (1970) Can we and should we study race differences? In J. Hellmuth (Ed.), *Disadvantaged child*, Vol. 3. *Compensatory education: A national debate*. New York, Brunner/Mazel. Pp. 124-157.
- Jensen, A. R. (1971) The race \times sex \times ability interaction. In Cancro, R. (Ed.), *Contributions to intelligence*. New York, Grune & Stratton. Pp. 107-161.
- Jensen, A. R. (1972a) The causes of twin differences in IQ. *Phi Delta Kappan*, 53, 420-421.
- Jensen, A. R. (1972b) Discussion of Tobach's paper. In Lee Ehrman, G. S. Omenn, and E. W. Caspari (Eds.), *Genetics, environment and behavior: Implications for educational policy*. New York, Academic Press.
- Jensen, A. R., (1972c) *Educability and group differences*. London, Methuen (New York, Harper and Row, 1973).
- Jensen, A. R. (1973) *Educability and group differences*. New York, Harper and Row (London, Methuen).
- Jinks, J. L., and Fulker, D. W. (1970) Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of human behavior. *Psychol. Bull.*, 73, 311-349.
- Leahy, Alice M. (1935) Nature-nurture and intelligence. *Gen. Psychol. Mono.*, 17, 241-305.
- Light, R. J., and Smith, P. V. (1969) Social allocation models of intelligence: A methodological inquiry. *Harv. educ. Rev.*, 39, 484-501.
- Light, R. J., and Smith, P. V. (1971) Statistical issues in social allocation models of intelligence: A review and a response. *Rev. educ. Res.*, 41, 351-367.
- Lush, J. L. (1968) Genetic unknowns and animal breeding a century after Mendel. *Transactions of the Kansas Academy of Science*, 71, 309-314.
- McNemar, Q. (1940) A critical examination of the University of Iowa studies of environmental influences upon the IQ. *Psychol. Bull.*, 37, 63-92.

- Noble, C. E. (1969) Race, reality, and experimental psychology. *Perspectives in biology and medicine*, 13, 10-30.
- Page, E. B. (1972) Miracle in Milwaukee: Raising the IQ. *Educ. Res.*, 1, 8-16.
- Scarr-Salapatek, S. (1971) Race, social class, and IQ. *Science*, 174, 1285-1295.
- Shockley, W. (1971a) Negro IQ deficit: Failure of a 'Malicious Coincidence' model warrants new research proposals. *Rev. educ. Res.*, 41, 227-48.
- Shockley, W. (1971b) Models, mathematics, and the moral obligation to diagnose the origin of Negro IQ deficits. *Rev. educ. Res.*, 41, 369-377.
- Skodak, Marie, and Skeels, H. M. (1949) A Final follow-up study of one hundred adopted children. *J. gen. Psychol.*, 75, 85-125.
- Spuhler, J. N., and Lindzey, G. (1967) Racial differences in behavior. In J. Hirsch (Ed.), *Behavior-genetic analysis*. New York, McGraw-Hill. Pp. 366-414.
- Terman, L. M. (1940) Personal reactions of the Yearbook Committee. In Whipple, G. M. (Ed.) *Intelligence: Its nature and nurture*, 39th Yearbook of the National Society for the Study of Education, Part I. Pp. 460-467.
- Tuddenham, R. D. (1970) A 'Piagetian' test of cognitive development. In Dockrell, B. (Ed.), *On intelligence*. Toronto, Ontario Institute for Studies in Education. Pp. 49-70.
- Vernon, P. E. (1965) Environmental handicaps and intellectual development: Part II and Part III. *Brit. J. educ. Psychol.*, 35, 1-22.