REPLIES TO SHOCKLEY, PAGE, AND JENSEN: THE CAUSES OF RACE DIFFERENCES IN I.Q.

By N. L. GAGE

This paper is intended as a reply to the comments by Shockley,1 Page,2 and Jensen3 on my recent paper,4 which opposed one by Shockley.5 I have organized this reply in terms of the basic issues raised in their comments: 1) heritability and improvability, 2) genetic research approaches, 3) environmental research approaches, 4) I.Q. and environment-difference correlations, and 5) the heritability of race differences.

Heritability and Improvability

It is true — as Shockley, Page, and Jensen indicate — that I did not refer specifically to the heritability of I.Q. The reason is that such heritability estimates are only tangentially related to the basic question, the one with which Jensen reopened the issue in 1969: “How much can we boost I.Q. and scholastic achievement?” I welcome Jensen’s assertion that the kind of analysis I emphasized “is a legitimate and worthwhile kind of analysis.” To my mind, it is far more central to the question of how much we can boost I.Q. and achievement than is the heritability index.

Heritability has been misinterpreted as meaning the opposite of improvability through environmental influence. But it is untrue that the higher the heritability index the less I.Q. can be changed by environmental influences. The heritability index of 80%, on which all three critics place such great emphasis, reflects only a certain set of conditions, both genetic and environmental. And the environmental part of those conditions can often be controlled by man. The conditions on which the 80% estimate is based are those that have operated at a given time in the past (say, the first two-thirds of the twentieth century), in a certain place (northern Europe and the United States), and on a given part of the population (whites). The 80% estimate tells nothing about what could happen in the future, in the United States, to blacks.

To clarify this point, let us consider a trait with even higher heritability than that estimated for I.Q. This trait is height, whose heritability is about 90%. If this high heritability index had been derived in the year 1800, would it have been safe to conclude that height cannot be increased through environmental influences? If that conclusion had been drawn, it would have been wrong:

For the last 170 years, people have been growing bigger and maturing earlier. In most countries, the average size of the adult has been increasing at the rate of one inch every 30 to 50 years . . . . These changes are based on improved nutrition, reduced energy expenditure, and better medical care. In all countries better nourished children grow taller and mature earlier. The level of nutrition has been raised considerably in most parts of the world by the elimination of vitamin deficiencies, the provision of more and higher quality protein, and the supplementation of caloric intake.6

Just as heritability was largely irrelevant in 1800 to the question of how much we could boost height, so it is similarly irrelevant in 1972 to the question of how much we can boost I.Q. and scholastic achievement. In my paper, I showed that the correlation between the I.Q.’s of separated identical twins, and hence the heritability index, is a function of the variance in environment. As was indicated by Bloom’s analysis, which I quoted, if only identical twins placed in highly similar environments are considered, the correlation between their I.Q.’s goes up; if only those placed in very different environments are considered, the correlation between their I.Q.’s goes down. The respective heritability indexes would also go up or down. The heritability index of 80% reflects the variances that happen to have occurred in the environments of identical twins reared apart and in those of persons in other blood and rearing relationships. These variances are functions of accidents of foster-child placement and of the distribution until now of I.Q.-relevant environments.

Most separated identical twins are placed in similar environments. This fact is indicated by the distribution of differences in I.Q. between separated identical twins, a distribution reported by Jensen.7 The shape of this distribution of differences in I.Q. is, as Jensen pointed out, approximately like that of one-half of the normal curve. It is sharply skewed to the right. The shape of the distribution of environment differences is probably similar, as is suggested by the distribution of ratings of estimated educational advantages in the data of Newman et al., shown in Figure 1. The median differences in estimated educational advantages in the data of Newman et al. have values of 12. The two cases that had these median values are Cases III (Paul C. and Paul O.) and XIV (Esther and Ethel). How different were their educational advantages?

Quantitatively, it appears that Paul C. has had about seven months more schooling than Paul

N. L. GAGE (2448, Stanford University Chapter) is professor of education and psychology at Stanford University and chairman of the Executive Board, Stanford Center for Research and Development in Teaching. He wishes to acknowledge helpful criticisms and suggestions made by Bruce J. Biddle, John A. Creager, Janet D. Elashoff, and Richard E. Snow.
O, and in larger, if not better, schools. The difference is not great. . . . The economic status of the two foster-families was much the same, Paul C.'s foster-father being a painter and paperhanger, and Paul O.'s a telegraph operator. There are few, if any, cultural influences in either home . . . and there have been no marked contrasts in the living conditions of the two homes. . . . Paul C. has had somewhat better educational advantages [he had just finished a three months' course in accounting at a business college, while Paul O. gave up engineering school before the first year was over].

Ethel lived in a Catholic orphanage from age 6 to 19 and received her entire education in connection with the orphanage. She had roughly a grade-school education. Esther finished the eighth grade in a country school and afterward took a selling and tailoring course.

The difference in quality of education is difficult to estimate, but Ethel's schooling extended through a larger number of years and probably for more months in each year. . . . On the whole, the social environmental milieus have been moderately similar but with some differences. . . . The contrast in education or other environmental influences appears to have favored Esther's intellectual performance, but the differences are relatively small.

All in all, it appears from these descriptions that the median pairs of separated identical twins in the data of Newman et al. had highly similar environments. Although it is a matter of subjective judgment, my own impression is that this median difference is much smaller than that which would be found between representative samples of black and white environments in the United States.

The shape of the environment-difference distribution is not foreordained by natural law; it is subject to human control. If the shape of the distribution of environmental differences were made just the opposite, so that most identical twins were placed in extremely different environments, falling into a distribution skewed to the left, the heritability index based on the correlation between their I.Q.'s would be much smaller.

The same point - that the heritability index depends on the man-made variance in environment - holds for other kinds of data on which heritability indices are based, i.e., for other blood and rearing relationships. Shockley himself implied that heritability indices are relative to environmental variances when he added the qualification "for individuals from families like those that raise one of a pair of white identical twins." It is highly unlikely that the I.Q.-relevant environmental differences in such families are as great as those between white and black families in the United States.

What kinds of environmental effects could be obtained through the application of psychological research that had pierced into the major causative factors in I.Q. development? The answer is as yet unknown. But human height also went unchanged for eons. It improved dramatically after nutritional science discovered such crucial factors as vitamins. Environmental influences will produce similarly substantial improvements in I.Q. when science finds the crucial environmental variables - and society uses those findings.

**Genetic Research Approaches**

Why did I not deal with Shockley's suggestions of a "variety of methods for research on genetic contributions to I.Q. deficits applicable to Negroes"? The papers to which Shockley referred (in his footnotes 18-20) appear to be as irrelevant to the main issue - the improbability of I.Q. and scholastic achievement - as the heritability index. Shockley's Figure 5, for example, apparently shows a correlation between percentage of Caucasian ancestry in blacks and percentage of rejection of blacks for mental reasons in various Army recruiting districts. Without information as to Shockley's "ways of controlling for the environmental differences," his proposed method cannot be evaluated. His Figure 6 dealt with visual acuity and not I.Q. or scholastic achievement. The bearing of the methodology on the improbability of I.Q. and achievement is remote at best. His Figure 7 dealt with the correlation coefficient between achievement variables and personality variables. Such coefficients do seem somewhat relevant to the improbability of I.Q. and achievement. The racial differences in correlation coefficients may indicate that present-day teaching methods affect the achievement of blacks differently than they affect the achievement of whites. Such an "aptitude-treatment interaction," if validated, would suggest searching for teaching methods more effective in maximizing the achievement of blacks. But it has nothing to do with the question of whether the black-white difference is genetic or environmental in origin.

Shockley says, "I differ most with the Gage position on the possibility of inventive approaches. . . ." I said nothing in my paper about this possibility; my reference to the "most important way to disprove Jensen's hypothesis" as that of reducing race differences "through education and other kinds of environmental influence" still stands. Genetic research along the merely correlational lines thus far proposed entails no improvements of the environment. Only experimental manipulations which change environmental variance and improve the educational and occupational
achievement of blacks can definitively test the hypothesis that "genetic factors are strongly implicated in the average Negro-white intelligence difference...". One tests a hypothesis by attempting to disprove it. Further correlational research cannot be as definitive as experimental research in disproving Jensen's hypothesis. As I also indicated, the social as well as the purely scientific issues require effort to improve the environment. Without that effort, neither the scientific nor the social problems can be solved.

Shockley notes the lack until now "of a widely recognized study of the conquest of black disadvantages by environmental improvements." This lack is equivalent to the lack thus far of cures for schizophrenia or cancer. Does Shockley advocate that the search for environmental or noneugenic ways of treating schizophrenia and cancer should therefore be abandoned? Is Shockley convinced, prior to any of the research for which he asks, that the geneticity of I.Q. makes educational efforts altogether unpromising?

Shockley's references to the research on genetic aspects of learning and memory in mice, by Bovet et al., seem to me to be altogether irrelevant to the issues at hand. What mice have to offer as foster parents has nothing to do with what human beings as educators have to offer human children. Indeed, human beings can train mice more effectively than other mice can train mice. This example does little to strengthen Shockley's allegation that such research suffers from policy taboos. If Shockley's opinions about the hereditary causation of Negro intellectual deficits are based on "considerations like these" (i.e., extrapolations from Bovet's mice studies), they have even less basis than when they are based on extrapolating white-twin geneticity data.

In setting research-support policies, we should reject any blanket endorsement of either genetically or environmentally oriented research. Instead, specific research proposals should be reviewed on their merits by technically competent referees. The proposals should be supported if, in the judgment of such referees, they have sufficient promise of scientific or social value.

Environmental Research Approaches

The difference between Shockley and me is indeed well typified by his section on "Quantifiable Humanism?" and mine on "Better Educational R & D Needed." In that section Shockley regards "sound diagnosis" of human quality problems as essential. He regards evidence as to the heritability of educational and occupational deficits of blacks as the goal of such sound diagnosis. Once the goal is achieved, we should take steps to minimize the birth-rate of persons with low I.Q.'s. (Such a policy, under present-day circumstances, would of course mean disproportionate reductions in the birth-rate of blacks.)

"To say that 'compensatory education has been tried,' with the implication that it should be abandoned, because billions of dollars have been spent on it, is a non sequitur. Because billions of dollars have been spent on unsuccessful cancer treatments, should we say that cancer treatments should no longer be sought?"

My own approach is to seek ways of improving the environment, especially the educational part of it. To say that "compensatory education has been tried," with the implication that it should be abandoned, because billions of dollars have been spent on it, is a non sequitur. Because billions of dollars have been spent on unsuccessful cancer treatments, should we say that cancer treatments should no longer be sought? Those billions of dollars for compensatory education were spent in ways inadequately based on educational research. Only a tiny fraction of that money was spent on research. It is true that the particular approaches to compensatory education that have been tried thus far have not yielded dramatically positive results. But we should not conclude that new approaches, based on better research, are no longer worth seeking. To do so is to violate lessons learned throughout the history of scientific research and development. For example, we could say that "the curing of cancer (or heart disease, or schizophrenia) has been tried and it apparently has failed" because millions of persons still cannot be cured of cancer (or heart disease, or schizophrenia). Past failures indicate the need for better research, not the cessation of the effort. Unfortunately, Jensen's statement was widely interpreted as justifying the abandonment of effort to improve education for socially disadvantaged segments of the population.

Shockley disagrees that such programs have continued for "merely a single enthusiastic Presidential administration." I meant, of course, that only in the Johnson administration were vast new programs launched. It is true that such efforts go back to "at least 1956." Thus, in 1972 compensatory education has a history of perhaps 16 years of serious effort. And it has not yet had any resounding successes! (In what other fields of science and technology has patience been so short?)

We should sincerely welcome Shockley's statement that he is "not anti-educational research." I urge that he publish technical reports on the efforts that have markedly improved the grade-point averages of Stanford undergraduates. It is also heartening that Shockley seems impressed with Heber's work, which improved compensatory education. If Heber's approach is too expensive now, its costs per pupil may become feasible when it is refined and undertaken on a large scale.

The meaning of Page's conception of "equalizing opportunity" is unclear. In this instance, I agree with Jensen that

A large part of the failure, I believe, has resulted from the failure... to take seriously the problems of individual differences in development rates, patterns of ability, and learning styles... and educators are now beginning to say: "Let's really look at individual differences and try to find a variety of instructional methods and differentiated programs that will accommodate these differences."11

Jensen and I may disagree as to how much these individual differences are due to environmental factors. But we and, I think, Page agree on the importance of taking them into account. If black children from low-income areas differ in pattern and level of learning abilities from white middle-class children, then equalizing opportunity requires different treatment of the chil-
"Clearly, any adequate program aimed at improving the I.Q. and scholastic achievement of blacks must have its prenatal, medical, nutritional, and other biological aspects as well as its postnatal, education, and cultural aspects."

I.Q.- and Environment-Difference r's

Shockley deals with my reference to Burt's study by furnishing the scatter-plot shown in his Figure 1.12 Here the correlation between occupational class advantage and I.Q. advantage of 35 identical twins reared apart is only .34. This coefficient is apparently to be contrasted with the value of .79 that I cited from Newman, Freeman, and Holzinger. Why is the value of .34 not higher? The answer is simply that occupational class is an inadequate index of the I.Q.-relevant aspects of the environment.

It is true that the co-occurrence of the two rare events—the Gladys-Helen I.Q. difference and the Gladys-Helen environmental difference—does not prove a causal connection between them. I was rejecting the impression of a merely chance occurrence created by Jensen's and Shockley's treatment of the Gladys-Helen difference in I.Q. Both had failed to mention that Gladys and Helen differed most not only in I.Q. but also in educational advantage. The question is not whether the Gladys-Helen difference is compatible with 80% heritability. Rather, does the Gladys-Helen conjunction of I.Q.-difference and environment-difference throw light on the possibility that the black-white difference in I.Q. is attributable to environmental differences?

My emphasis on the correlation of .79 was not aimed at refuting the .80 heritability estimate. Rather, it was aimed at indicating the dependence of that estimate on the distribution of man-made differences in environment. The correlation of .79 implies that the reduction of race differences in I.Q.-relevant environment will reduce race differences in I.Q. The coefficient of .79 implies that such I.Q.-relevant factors can be found in the prenatal and postnatal environment. Thus, I must reply to Professor Shockley with a counter-challenge: Does the 80% heritability estimate preclude the substantial reduction of race differences in I.Q. and achievement through improvements in education and other environmental factors?

Jensen notes that Newman et al. used three other intelligence tests besides the Stanford-Binet—and that twin differences on these other intelligence tests correlated .55, .57, and .46 with educational-advantage differences. These correlations are far from being so low as to lack great educational significance. These data (like Burt's) substantially indicate that I.Q. differences are associated with educational-advantage differences even when genetic differences are nonexistent.

As Jensen notes, the I.Q. differences of separated identical twins are correlated with their fingerprint differences. Much other evidence13 also indicates that prenatal nongenetic factors, in addition to educational-environmental differences, may have caused some of the I.Q. differences between separated identical twins. This possibility is supported by findings on relationships of I.Q. to prenatal environment, birthrate, and reproductive casualties (owing to pregnancies at early ages, in close succession, in large numbers, or late in reproductive life). Such I.Q.-lowering prenatal environmental influences occur much more frequently among poor women and blacks. Thus, for example, "Prematurity and low birth-weight have a markedly higher incidence among Negroes than among whites."14 Clearly, any adequate program aimed at improving the I.Q. and scholastic achievement of blacks must have its prenatal, medical, nutritional, and other biological aspects as well as its postnatal, educational, and cultural aspects.

Jensen states that he did not choose the lower of two correlations between I.Q.- and environment-difference in his discussion of Burt's study, i.e., the correlation for the individual intelligence test. But then he proceeds to explain why he did choose the lower of the two correlations. Perhaps he means that he did not unjustifiably choose the lower one. It is good to have his explanation that he wanted to use Burt's data based on individual tests because the other three studies of separated identical twins had also used individual rather than group tests. The issue here is one of judgment in reviewing and summarizing research. I would have preferred to see much fuller detail on these important matters, e.g., the correlations for all of the tests used in the studies by both Newman et al. and Burt. And I should have based my judgment as to the comparability of the tests not so much on whether they were individually or group administered as on their content and their correlations with other tests. In any event, the median correlation for the group tests used by Newman et al. is .55. The value of r for the group test used by Burt is .43.

Jensen further justified his choice of the data on individually administered tests on the grounds they were not subject to spuriously inflated correla-
Jensen explains Burt's "final assessment" in an attempt to show that I misunderstood its intent. But then Jensen states that Burt's intention was indeed "to yield a higher heritability estimate than the raw scores on a single I.Q. test." Burt should, of course, have been intending to maximize the validity of the heritability index, not its magnitude. Burt submitted the I.Q. scores of the children to the teachers for comparison with the "teachers' judgments of their pupils native intelligence" (italics mine). The teachers either knew or were informed that Burt defined intelligence as an innate (i.e., inherited) characteristic. So they were perhaps unwittingly influenced to exhibit exactly the kind of bias with which I am concerned. Is it not likely that only those children were retested with two or more other I.Q. tests whose initial I.Q. scores were out of line with Burt's or the teachers' knowledge of the I.Q. of the real parents or the other twin? To work on the scores in order to maximize heritability is, it seems to me, to commit the logical fallacy of begging the question. That is, Burt and Jensen demonstrated their conclusion by means of premises which presuppose that conclusion. Thus, they seem to be seeking to demonstrate high heritability by means of measures of intelligence designed to maximize heritability. My suspicions concerning the low correlation of .15 between I.Q. differences and environment differences, when I.Q.'s were based on "final assessments," seem to me to be warranted.

Heritability of Race Differences

Jensen asserts that:

High heritability of a trait within populations that differ in the trait does, however, increase the a priori likelihood of a genetic difference between the populations. The fact of the high heritability of I.Q., therefore, makes it a very reasonable and likely hypothesis that genetic factors are involved in the Negro-white difference. No geneticist to my knowledge has argued otherwise.

I find it difficult to understand how Jensen could write the last sentence, above, in view of his debate with . . . make it highly improbable that the environmental influences which contribute to the environmental variance of I.Q. in twin studies are anywhere near sufficient to account for a 15-point I.Q. difference between two populations. . . . There would have to be practically no overlap (i.e., a difference of 3.5 to 4.5 standard deviations) between the Negro and white distributions of quality of the environment (if by environment we mean those factors which make for a difference between MZ twins reared apart) to account for a 15-point I.Q. difference. So far no one has hypothesized in a testable fashion any other nongenetic factors that could explain this difference.

But, as he noted in an earlier paper, his scale of effects of environmental postnatal factors. These small values make it highly improbable that the environmental influences which contribute to the environmental variance of I.Q. in twin studies are anywhere near sufficient to account for a 15-point I.Q. difference between two populations. . . . There would have to be practically no overlap (i.e., a difference of 3.5 to 4.5 standard deviations) between the Negro and white distributions of quality of the environment (if by environment we mean those factors which make for a difference between MZ twins reared apart) to account for a 15-point I.Q. difference. So far no one has hypothesized in a testable fashion any other nongenetic factors that could explain this difference.

Jensen noted the small values of the mean and standard deviation of the I.Q. differences between separated identical twins that can be attributed to environmental postnatal factors. These small values make it highly improbable that the environmental influences which contribute to the environmental variance of I.Q. in twin studies are anywhere near sufficient to account for a 15-point I.Q. difference between two populations. . . . There would have to be practically no overlap (i.e., a difference of 3.5 to 4.5 standard deviations) between the Negro and white distributions of quality of the environment (if by environment we mean those factors which make for a difference between MZ twins reared apart) to account for a 15-point I.Q. difference. So far no one has hypothesized in a testable fashion any other nongenetic factors that could explain this difference.

As in his original article, Jensen in his reply relies heavily on the weight of authority as relevant evidence. We hear of a "Nobel laureate in physics," or "three sociologists who are students of this problem" and who "all agree," "geneticists such as K. Mather, C. D. Darlington, R. A. Fisher, and Francis Crick, to name a few," and finally, "a number of highly qualified geneticists" who have reviewed this "treatment of quantitative genetics and have found no fault with it." Well, I am a very highly qualified geneticist whose field is the study of genetic variation in natural populations, and I found a few faults.

Lewontin: "Well, I am a very highly qualified geneticist whose field is the study of genetic variation in natural populations, and I found a few faults "[with Jensen]."

Richard Lewontin. Lewontin had stated that:

The fundamental error of Jensen's argument is to confuse heritability of a character within a population with heritability of the difference between two populations. Indeed, between two populations, the heritability of their difference is meaningless. This is because a variance based upon two measurements has only one degree of freedom and so cannot be partitioned into genetic and environmental components. The genetic basis of a difference between two populations bears no logical or empirical relation to the heritability within populations and cannot be inferred from it. . . .

Then Lewontin goes on to illustrate how two completely inbred lines of corn, with no genetic variance within each group, will show between-group variance in height due to genetic differences in the two lines. Here the heritability within groups is zero, but the two lines differ in mean height nevertheless.

In the opposite condition, two varieties of corn are used that have much genetic variance between varieties. Here the heritability of their difference is 100% heritability or 100% heritability. The heritability within groups tells nothing about the heritability of the difference between groups. The reader is urged to review the entire Lewontin-Jensen exchange11, 17, 18 to appreciate the cogency of this analysis.

Now, who is Lewontin? He is a professor of biology at the University of Chicago. In his final article in the exchange with Jensen, he says:

All use subject to http://about.jstor.org/terms

This content downloaded from 128.95.155.210 on Fri, 22 Sep 2017 23:54:52 UTC
All use subject to http://about.jstor.org/terms
effects of environmental differences is probably much smaller than what would be obtained if we paired randomly chosen black and white environments. With a larger value for the correct estimate, the difference between black and white distributions of quality of environment would not need to be as great as "3.5 to 4.5 standard deviations" to account for the 15-point I.Q. differences. The necessary difference is probably much more within the empirically probable limits. Thus an environmental explanation of the I.Q. difference remains plausible.

In my comments on Shockley's "thinking exercise" - the "voluntary sterilization plan" - I raised questions concerning the regression effect, the low predictability of children's I.Q. from those of their parents, and the idea of making I.Q. the sole nonphysical criterion of desirability of parenthood for nonpayers of income tax. Shockley answered none of these questions.

In addition, of course, any such thinking exercise should provoke the question of its financial cost. Shockley is concerned about the money needed for Heber's methods of obtaining marked increases in I.Q. But he does not consider such costs in presenting his sterilization bonus plan.

Shockley's reference to the high I.Q.'s of children of Terman's gifted group is irrelevant; it does not belte the well-established correlation of only .5 between parent I.Q. and children's I.Q. Nor does it belie the established fact that parents with I.Q.'s below 100 will, on the average, have children with higher I.Q.'s. In his next-to-last sentence, Jensen asks for clear and testable hypotheses. Fortunately, behavioral scientists have not waited for Jensen to set this task. And Jensen himself, with his hypothesis concerning improved methods of educating children low (for whatever reason) in abstract intelligence, has participated in the search. As the search pays off, we shall answer the question of how much we can boost I.Q. and scholastic achievement. The answer should not be reached by default, through passive acceptance of a status quo buttressed by heritability estimates. To do so would be equivalent to accepting schizophrenia, cancer, or diabetes as essentially untreatable because they also have their hereditary components.

Finally, let me note that none of my critics has commented on the point I made about the very high correlation between school attainment differences and environmental differences for identi-