A Reply to Gage: The Causes of Twin Differences in I.Q.
Author(s): Arthur R. Jensen
Source: *The Phi Delta Kappan*, Vol. 53, No. 7 (Mar., 1972), pp. 420-421, 419
Published by: Phi Delta Kappa International
Stable URL: http://www.jstor.org/stable/20373252
Accessed: 22-09-2017 23:56 UTC

A REPLY TO GAGE:

THE CAUSES OF TWIN DIFFERENCES IN I.Q.

BY ARTHUR R. JENSEN

Gage's discussion of I.Q. differences between identical or monozygotic (MZ) twins is likely to be very misleading to those who are not thoroughly knowledgeable in this area of research.[1]

Although Gage's readers may have gained the impression that he has brought into question the estimates of the heritability of intelligence as derived from twin studies, this is not the case at all. The most reliable estimates of the heritability of I.Q. (that is, the proportion of individual differences variance attributable to genetic factors) yield values in the range from about .7 to .8. Gage has presented nothing that in the least questions this conclusion based on published studies of the heritability of I.Q. Instead, what he has done is to discuss the estimated proportional contributions of various environmental factors to the nongenetic part of the total I.Q. variance, that is, the remaining 20 to 30% after the genetic variance (of 70 to 80%) is removed. Well and good. This is a legitimate and worthwhile kind of analysis. But it must be clearly understood for what it is.

I.Q. differences between MZ twins (since they are genetically identical) can reflect only nongenetic influences. Gage places great emphasis on the finding in the study by Newman, Freeman, and Holzinger[2] (hereafter abbreviated NFH) of a correlation of .79 between Stanford-Binet I.Q. differences in their 19 pairs of MZ twins reared apart and the rated differences in their educational advantages. In other words, .79 or 62.4% of the *non*genetic portion of the variance in I.Q.'s is associated with the rated differences in educational advantages between the separated twins.

ARTHUR R. JENSEN is professor of educational psychology at the University of California, Berkeley, and author of numerous works on the heritability of intelligence.

Now, let's take a closer look at this. The correlation of .79 looks quite large. Indeed, it is large, as correlations go. It is the highest correlation between educational differences and I.Q. differences to be found in the whole literature. The highest correlation of this type found by Burt[3] on a much larger sample of twins is only .43. Gage does not mention the fact that NFH used three other intelligence tests besides the Stanford-Binet. Twin differences on these other intelligence tests correlated with educational differences .55, .57, and .46 – all much lower than the .79 for the Stanford-Binet I.Q.

But let's stay with the highest correlation of .79 and see what it actually means. The average difference in I.Q.'s between the 19 sets of twins was 8.21 I.Q. points. Assuming a Stanford-Binet reliability of .95, or 5% error variance in scores, the rated educational differences between the twins accounts for 5.43 points of the I.Q. difference. But are these 5.43 I.Q. points all really *caused* by the *educational* differences between the twins? Very unlikely. Here is the evidence: NFH also measured the degree of similarity of fingerprints (ridge count) of their 19 sets of twins. Any difference in fingerprints is also nongenetic, of course. But the "environmental" influences that make for fingerprint differences in MZ twins must have occurred before the third or fourth month of fetal life; fingerprints do not change (except in overall size) beyond this point in prenatal development.[4] Yet in the NFH study the fingerprint differences correlate .51 with Stanford-Binet I.Q. differences and .48 with educational differences (both correlations significant beyond the 5% level). This can only mean that some of the .79 correlation between I.Q. difference and educational difference is attributable to the association of each with a third factor – something in early intrauterine development. These early prenatal effects are much greater in identical twins than in fraternal twins and singletons, as indicated by the higher incidence of congenital malformations, fetal loss, and disparity in birth weights of MZ twins.[5] So if the common association with fingerprint differences is taken account of, the differences in educational advantages account for something less than 5 I.Q. points. The famous case of Gladys and Helen, the twins who differed by 24 Stanford-Binet I.Q. points (the largest MZ twin difference ever found), is especially interesting because they also show the largest educational difference (third-grade education *vs.* college graduate) AND *the largest difference in fingerprints.* They differ by 4.13 standard deviations in fingerprint ridge count; in this measure they are much more like fraternal than like identical twins. It is most probable that a large part of the I.Q. difference between Gladys and Helen is attributable to factors occurring very early in gestation rather than to factors of a cultural, psychological, and educational nature in their postnatal environments, as extremely different as these were. Also, it should be noted that one such extreme case in a sample of only 19 twin pairs can significantly affect the average statistics. That is why I give more weight to Burt's study,[6] with 53 twin pairs, and have reanalyzed the combined original data of the four largest studies of the I.Q.'s of MZ twins reared apart, totaling 122 pairs.[7] In that article I cite evidence that differences in birth-weight of twins are correlated with later I.Q. differences; also, it has been found that twin differences in height and other skeletal measurements are significantly correlated with I.Q. differences.[8] These are not cultural-psychological effects, but biological and prenatal effects.[9] It is more correct to call them "nongenetic" rather than "environmental," since most persons think of "environment" as the postnatal social-cultural-educational environments which the "environmentalists" keep telling us are such important causes of the great spread of individual differences in mental abilities that we observe in schools, in the armed forces, and in the world of work.*

Gage[10] accuses me of choosing the lower of two correlations in my discussion[11] of Burt's study, i.e., the correlation for the individual intelligence test. This is exactly what I did *not* do. Since I was combining the original data of four major twin studies, three of which were based entirely on individually administered tests, it was only methodologically correct that I should use Burt's data based on individual tests (the English adaptation of the Stanford-Binet) rather than the group-administered tests which were thereby not as comparable to the tests used in the other three studies. In short, for good methodological reasons I chose the *tests* and *not* the correlations they may have yielded with other variables.

Gage is a bit amiss, too, in claiming that individually administered tests are not generally superior to group administered tests for this type of research, based as it is upon correlations. Unreliability attenuates correlations and the administration of tests in classroom groups can spuriously inflate correlations if twins are in the same classes or if two or more intercorrelated tests were administered by the same teacher under similar biasing conditions. When children are tested in classroom groups by their teachers, any biasing influence on test performance, such as poorly given instructions and lax observance of the time limits for standardized tests, contributes to "between-teacher" variance which adds a spurious increment to the intraclass correlations involving these tests. So there is good reason to prefer individual tests for correlational purposes; and if group tests are used they should be given by specially trained testers, not the classroom teachers, and when two or more tests are used they ideally should be administered by different testers.

Gage also seems to misunderstand the intent of Burt's "final assessment" scores. Gage writes: ". . . the 'final assessment' could readily, even if unintentionally, have been biased in such a way as to reduce its tendency to reflect a child's environment and increase its conformity to the child's hereditary background."[12] This makes it look as though the adjusted assessments might have been "fudged" to yield a higher heritability estimate than the raw scores on a single I.Q. test. But this was precisely Burt's intention. I discussed it with him personally while in London last summer. What he tried to do was to determine whether teachers' judgments of their pupils' native intelligence, when used along with test scores, could increase the heritability of these teacher-adjusted assessments. Burt submitted the I.Q. scores of all the children in the class to the teacher for examination and criticism. Teachers were told to note those scores which under- or over-estimated their own judgment of the child's ability. In these discrepant cases, the child was retested on two or more other I.Q. tests and the composite scores were used. The effect of the adjusted assessment was to increase the reliability and validity of the group-administered test scores as indicators of innate intelligence. The fact that this was possible is shown in the marked decrease (10.6 to 1.43%) in the between-families component of environmental variance and in the covariance of genetic and environmental factors (which Burt combines under the label "systematic environmental effects") in the adjusted assessments as compared with the unadjusted test scores.[13] (Note that other sources of variance were only slightly affected by the adjusted assessments.) If the adjusted assessments were more influenced by the child's socioeconomic or cultural background than the unadjusted scores, the effect on the magnitude of these components of environmental variance should be just the opposite to what was actually found.

Finally, what is the relevance of these studies to determining the causes of the Negro-white I.Q. difference which in the United States today averages about one standard deviation (i.e., 15 I.Q. points)? Three main points, I believe, would receive the assent of geneticists who have studied the matter. First, twin correlations and the estimation of the heritability of I.Q. *within* each of two populations does not provide any formal proof that an observed average difference between the populations is attributable to genetic factors. Other genetical methods than heritability analysis are required to determine the extent of genetic difference between two populations. (For more detailed discussion of these points the reader is referred elsewhere.)[14-16] Second, 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 I.Q. difference. No geneticist to my knowledge has argued otherwise. Third, the small values of the mean and standard deviation of the distribution of the I.Q. differences between MZ twins reared apart that can be attributed to environmental factors, particularly postnatal factors, 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. As I have pointed out elsewhere,[17] 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 differences 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. Our scientific understanding of this problem can be advanced only by formulating clear and testable hypotheses. Hortatory rhetoric, however nobly motivated, will get us nowhere.

---

*Some geneticists, such as the cytogeneticist C. D. Darlington ("Heredity and Environment," *Proceedings of the 9th International Congress of Genetics,* 1954, pp. 370-81), maintain that identical twin differences overestimate environmental effects, since some of the difference is due to unequal division of the fertilized ovum, creating what Darlington terms "cytoplasmic discordances" and "asymmetry," which result in inequalities beginning at the earliest stages of development. There are also inequalities in blood supply – a condition peculiar to MZ twins. Newman, Freeman, and Holzinger note these phenomena in their discussion of fingerprint differences between twins (p. 119). NFH state, ". . . it may be regarded as proven that a considerable amount of asymmetry reversal, rarely complete and only slight in extent, occurs as a concomitant of monozygotic twinning. Asymmetry reversal, especially partial asymmetry reversal, causes differences in identical twins. Such differences are neither genetic, in the ordinary sense, nor environmentally induced. In comparing the variability of identical and fraternal twins, therefore, it is not proper to consider all differences in identical twins reared together as environmentally determined" (p. 51).

---

1. N. L. Gage, "IQ Heritability, Race Differences, and Educational Research," *Phi Delta Kappan*, January, 1972, pp. 308-12.

2. H. H. Newman, F. N. Freeman, and K. J.

(Continued on page 419)

learners are fostered by the emotionally stable good learners? No! On the contrary, their freezing responses increase so much – about 10% to 50% – as to reduce their success at shock avoidance. Good learners do not become freezers when reared in the home environment of freezing foster parents. Their avoidance scores do drop, however, but only slightly – about 80% instead of 87% on the fourth day.

Would research on the behavior genetics of different subgroups or races of Homo sapiens establish similar effects? Do Bovet's mice studies indeed suggest the reason for the lack of reports of statistically significant successes in transracial adoptions of black slum orphans?

My previously stated opinion that the major causes of the American Negro's intellectual and social deficits are primarily hereditary and racially genetic in origin has been reached as a result of considerations like these. The presentation in Gage that such conclusions are based on extrapolating white-twin geneticity data is an error that may permit the intellectual community to avoid the moral obligation to think. I have tried to philosophize on this matter of moral obligation. The results are available in written form as "Three Moral Postulates: Truth – Concern – Death."[12]

The KAPPAN-Voltaire Parallel

"I disapprove of what you say, but I will defend to the death your right to say it." So is Voltaire quoted.

My experience with the *Kappan* has parallels with both of Voltaire's clauses. The *Kappan* has indeed created and defended my right to express what I think and has done so knowing of the criticism it may face as a consequence. I have not always had my right to speak or be printed so zealously defended.

Once in each of the years 1968, 1969, and 1971, after a disruption or the threat of one prevented me from delivering a scheduled speech, Voltaire has been eloquently and publicly echoed – but to no effect. The president of Brown University's chapter of Sigma Xi, the honorary scientific research fraternity, complained bitterly about the cancellation of the twenty-fifth anniversary convocation of Sigma Xi at the Polytechnic Institute of Brooklyn, but my requests to speak either at Brown or P.I.B. were rejected. At Dartmouth College in 1969 my experience was similar. I am now awaiting a response from Sacramento State College (the case mentioned in the *Kappan* introduction) to my offer to lecture or participate in a TV debate with Professor Mercer or Mr. Mayeske; a spokesman for the Department of Health, Education, and Welfare has credited them with "refuting Jensen and Shockley." The *Kappan* lived up to Voltaire's second clause.

The *Kappan* also lived up to Voltaire's first clause – unwittingly, entertainingly, but, from my viewpoint, disappointingly. Consider this abbreviated quotation from the editorial introduction to the Shockley-Gage encounter:

"The editors . . . have no intention . . . of taking sides on the substantive questions. . . .

"We believe that bad conditions make bad people. We prefer to regard genetic inheritance . . . as simply . . . environment . . . and ultimately manipulable."

Is this not a claim of impartial non-side-taking promptly followed by rejection of my dysgenic threat? Indeed, if Bovet's mice are like people, would not the *Kappan's* environmentalist emphasis erroneously lead to the conclusion that "bad" emotionally unstable mice that freeze are made so by emotionally "bad" home environments? I propose that a striking inconsistency has arisen because the *Kappan*, in accord with Voltaire's first clause, strongly disapproves of what I say – indeed disapproves so strongly that what I say is rejected without the rejection being realized. If my analysis is sound, and this point gets across, then an inconsistency will have contributed to the principal objective of my strenuous writing efforts for the *Kappan* – to strengthen the moral obligation of my readers to think – even about dysgenics.

---

1. "Crisis in American Education," an address by Roger A. Freeman, special assistant to the President of the United States, before the annual meeting of the Washington State Research Council, June 19, 1970.
2. From Freeman manuscript printed in Extension of Remarks, *Congressional Record,* April 24, 1969, pp. E 3374-81.
3. Michael J. Wargo, Peggie L. Campeau, and G. Kasten Tallmadge, "Further Examination of Programs for Educating Disadvantaged Children," Final Report of Contract OEC-0-70-5016, Report No. AIR-2026-7/71-FR, July, 1971.
4. "Better Education for Minority Groups?," *Daily Gazette,* Berkeley, April 15, 1970, p. 9.
5. Cyril Burt, "The Genetic Determination of Differences in Intelligence: A Study of Monozygotic Twins Reared Together and Apart," *British Journal of Psychology,* 1966, pp. 137-53.
6. Lewis M. Terman and Melita H. Oden, "The Gifted Group at Mid-Life: Thirty-five Years' Follow-up of the Superior Child," *Genetic Studies of Genius* (Vol. 5). Stanford: Stanford University Press, 1959.
7. Edward Zigler, "Familial Retardation: A Continuing Dilemma," *Science,* January 20, 1967, pp. 292-98.
8. W. A. Kennedy, V. Van de Riet, and J. C. White, Jr., "A Normative Sample of Intelligence and Achievement of Negro Elementary School Children in Southeastern United States," Monograph 90, The Society for Research in Child Development, 1963.
9. As quoted from W. A. Kennedy, unpublished manuscript titled "Racial Differences in Intelligence: Still an Open Question?," in W. Shockley letter to John W. Gardner, printed in Extension of Remarks, *Congressional Record,* August 12, 1969, p. E 6849.
10. W. Shockley, "Human-Quality Problems and Research Taboos," *New Concepts and Directions in Education.* Greenwich, Conn.: Educational Records Bureau, 1969.
11. D. Bovet, F. Bovet-Nitti, and A. Oliverio, "Genetic Aspects of Learning and Memory in Mice," *Science,* 1969, pp. 139-49.
12. W. Shockley, "Three Moral Postulates: Truth – Concern – Death," *Letters, Presbyterian Life,* February 1, 1972.
13. W. Shockley, "A 'Try Simplest Cases' Approach to the Heredity-Poverty-Crime Problem," Proceedings of the National Academy of Sciences, June, 1967, pp. 1767-74. □

Jensen – footnotes

(Continued from page 421)

Holzinger, *Twins: A Study of Heredity and Environment.* Chicago: University of Chicago Press, 1937.

3. Cyril Burt, "The Genetic Determination of Differences in Intelligence: A Study of Monozygotic Twins Reared Together and Apart," *British Journal of Psychology,* 1966, pp. 137-53.
4. M. G. Bulmer, *The Biology of Twinning.* Oxford: Clarendon Press, 1970, p. 146.
5. *Ibid.*, p. 39.
6. Burt, *op. cit.*
7. Arthur R. Jensen, "I.Q.'s of Identical Twins Reared Apart," *Behavior Genetics,* 1970, pp. 133-46.
8. Barbara S. Burks, "Mental and Physical Development Patterns of Identical Twins in Relation to Organismic Growth Theory," *Yearbook of the National Society for the Study of Education,* Part II, 1940, pp. 85-96.
9. Burt, *op. cit.*, p. 142.
10. Gage, *op. cit.*, p. 309.
11. Jensen, *op. cit.*
12. Gage, *op. cit.*, p. 309.
13. Cyril Burt, "The Inheritance of Mental Ability," *American Psychologist,* 1958, pp. 1-15, Table 2.
14. Arthur R. Jensen, "Race and the Genetics of Intelligence: A Reply to Lewontin," *Bulletin of the Atomic Scientists,* May, 1970, pp. 17-23.
15. Arthur R. Jensen, "Can We and Should We Study Race Differences?," *Disadvantaged Child,* Vol. 3, J. Hellmuth (ed.). New York: Brunner/Mazel, 1970, pp. 124-57.
16. H. J. Eysenck, *The I.Q. Argument.* Freeport, N.Y.: The Library Press, 1971.
17. Arthur R. Jensen, "Twin Differences and Race Differences in I.Q.: A Reply to Burgess and Jahoda," *Bulletin of the British Psychological Society,* 1971, pp. 195-98. □