After each of the above trials, the recruiter in the aspirator was returned to the raid column. The effect of replacing her was dramatic; in one trial where only three workers had been recruited without the recruiter, 15 workers were recruited within 15 seconds, and by the end of the trial 30 workers had been recruited. Two tests that were unsuccessful without the recruiter were successful when she was replaced (Table 1).

These experiments show that the recruitment trail contains the essential information necessary for recruitment, but the response is lower than when combined with recruiter activity.

Response of workers to recruiter contact in the absence of a recruitment trail was determined by allowing potential recruiters to drop off a wasp nest to vegetation a few centimeters from a raid column. Ants on the column became visibly excited within seconds, and in a few minutes had ascended vegetation and were randomly searching upward in the vicinity of the wasp nest. The ants searched for almost a half hour, and several workers came within a few centimeters of the nest. Most likely, the ants would have located and attacked the nest except for the intervention of a hard rain. Similar searching has been observed prior to raids on other wasp nests. Thus, even without a recruitment trail, a recruiter releases searching behavior which may lead to prey capture.

Ant recruitment is based primarily on chemical and tactile signals. The expression and information content of those signals vary among ant species and result in diverse recruitment systems. Compared to other recruitment systems, that of army ants is among the most efficient in terms of gathering large numbers of workers quickly. In the relatively primitive system, called "tandem running," constant recruiter contact is necessary for orientation, and only one worker is recruited at a time (*14*). More advanced recruitment systems rely increasingly on chemical trails for orientation. In the system we term "group recruitment," orientation by the recruiter is still essential, but a chemical trail allows a large group of ants to follow a single leader. If the leader is removed, the group disbands (*3*). Group recruitment grades into "mass recruitment" in which the chemical trail is the primary orientator. Here the greatest efficiency is achieved. Less dependency on a single recruiter for orientation results in more continuous recruitment. The recruiter, however, still retains an important role. For example, in *Monomorium venustum* contact with the recruiter mobilizes the ants to search for an odor trail put down by the recruiter outside the nest. The ants do not follow the trail unless previously activated by the recruiter (*4*). In *Solenopsis invicta* the recruiter "alerts" workers to a chemical trail by vibratory antennal contact; the trail itself attracts and orients the recruits (*5*). Similarly, in army ants the recruiter enhances the response of workers to the recruitment trail.

In addition to an efficient recruitment mechanism, the most important feature of army ant foraging is the continuous raid columns that reduce delay between prey encounter and recruitment even 100 m or more from the bivouac. A comparison of initial recruitment rates for army ants and other ants indicates that army-ant recruitment is the highest (Fig. 1) (*4–10*). The comparison is made with some reservation since no standard procedure for measuring recruitment was used in the various studies. The results, however, agree well with expectation; the tandem-running ant, *Camponotus sericeus*, is slowest; the group-recruiter, *C. socius*, is intermediate; and the mass-recruiting *Solenopsis invicta* and *Eciton hamatum* are the fastest.

The adaptive value of rapid mass recruitment is shown by test raids of army ants on wasp nests. Small numbers of ants were readily thrown off nests by wasps, whereas large numbers caused wasps to abandon the nest, leaving their brood behind. The most common prey of army ants are colonies of insects and arthropods often larger than the ants themselves. Such prey can only be subdued by a massive and swift attack requiring efficient communication. The combination of continuous foraging columns, a recruitment trail that attracts and orients workers, secondary recruitment, and persistent recruiters results in the efficient gathering of the large attack force essential for army-ant raiding.

RUTH CHADAB
CARL W. RETTENMEYER
Biological Sciences Group,
University of Connecticut,
Storrs 06268

References and Notes

1. E. O. Wilson, *The Insect Societies* (Belknap, Cambridge, Mass., 1971), pp. 247–262.
2. M. G. Naumann, in preparation.
3. A. Collart, *Rev. Zool. Afr.* **15**, 249 (1927); C. Baroni Urbani, *Proceedings of the Seventh Congress of the International Union for the Study of Social Insects, London. 1973*, p. 12
4. R. Szlep and T. Jacobi, *Ins. Soc.* **14**, 25 (1967).
5. E. O. Wilson, *Anim. Behav.* **10**, 139 (1962).
6. B. Hölldobler, M. Möglich, U. Maschwitz, *J. Comp. Physiol.* **90**, 108 (1974), figure 2b.
7. R. H. Leuthold, *Psyche* **75**, 343 (1968).
8. B. Hölldobler, *Z. Vergl. Physiol.* **75**, 133 (1971).
9. ——— and E. O. Wilson, *Psyche* **77**, 393 (1970), figure 6a.
10. D. I. Wallis, *Behaviour* **23**, 167 (1964).
11. G. L. Ayre, *Can. Entomol.* **101**, 120 (1969).
12. M. S. Blum and C. A. Portocarrero, *Ann. Entomol. Soc. Am.* **57**, 793 (1964); J. F. Watkins II, *J. Kans. Entomol. Soc.* **37**, 22 (1964).
13. The aspirator consisted of a glass tube attached to a rubber tube with a piece of fine netting inserted between the tubes to stop insects.
14. E. O. Wilson, *Psyche* **66**, 29 (1959).
15. We thank L. Morales and M. E. Rettenmeyer for assistance during fieldwork, the Summer Institute of Linguistics for providing facilities at Limoncocha, Ecuador, and C. W. Schaefer and J. A. Slater for comments on the manuscript. Supported by NSF grant GB-31716 and the University of Connecticut Research Foundation.

31 January 1975; revised 11 March 1975

Heritability of IQ: Methodological Questions

The major thesis of Layzer's article "Heritability analyses of IQ scores: Science or numerology?" may be educed from his conclusion (*1*, p. 1265) that "published analyses of IQ data provide no support whatever for [the] thesis that inequalities in cognitive performance are due largely to genetic differences." From this two corollaries follow, one being that therefore no valid inferences can be drawn in respect to genetic differences in IQ between races, the other being that it is therefore pointless to speculate about the possible emergence of hereditary meritocracies. I will not deal with these corollaries but will attempt to examine the basic argument on which they rest.

This argument reduces to three main criticisms: (i) that the heritability concept is confused and the estimation of heritability (h^2) is feasible only if a number of unrealistic simplifying assumptions are made; (ii) that IQ tests have neither validity nor reliability; (iii) that, apart from the intrinsic defects of IQ tests, most of the data purporting to demonstrate that they measure differences which have some genetic basis are seriously flawed.

It is certainly true that the concept of heritability is complex and estimating it difficult. It seems to me, however, that Layzer compounds and exaggerates this complexity and difficulty. His figure 1, for example, shows the phenotypic responses of three genotypes, x_1, x_2, x_3, to a changing environment, y (in the caption to the figure the x and y are erroneously transposed). Now I find nothing particularly damaging to the heritability concept emerging from the hypothesized relationships depicted. Thus whether h^2 at y_1 should be greater or smaller than at y_2, where the development of the trait is maximal, constitutes an interesting and potentially soluble problem and not some basic and intractable mystery about heritability. Such problems are of great importance in behavior genetics and have, in fact, been considered by a great many researchers, notably Broadhurst and Jinks, Cattell, McClearn, DeFries, and Henderson.

Layzer approaches the analysis of heritability with the same attitude of hopelessness. Again, I do not see that his statistical formulations (*1*, p. 1260) demonstrate anything more than the proposition that genetic (G) and environmental (E) variables can interact and covary in ways perhaps contemplated but never specified by Ronald Fisher in his 1918 paper. Perhaps Fisher's definitions of G and E did entail "remarkable simplifications" (*1*, p. 1261), but their remarkableness is most certainly matched by that of Layzer's complications. Indeed, if he wishes (as implied by the title of his article) to champion "science" and to eschew "numerology," he might grant that in the initial stages of a genetic analysis it might be better to oversimplify than to overcomplicate.

Layzer insists (several times) that IQ scores contain "uncontrollable, systematic errors of unknown magnitude." It is difficult to fathom what he means by this phrase, since he uses it mainly to indicate, in blanket fashion, that measurement in the behavioral sciences is inferior to measurement in the physical and biological sciences. This may be true, but it should be pointed out that psychologists have dealt intensively with the fundamental problems of reliability and validity for over 70 years. The single reference Layzer cites for this work is a 1946 paper by S. S. Stevens which will hardly provide much illumination to readers of *Science*.

In point of fact, IQ tests are, on the average, highly reliable. The occasional large change in an individual's IQ is not the rule. It is the exception. Further, on the average, IQ scores do relate to a number of important external criteria, including level of education, occupation, social class, and most of all, chronological age. Thus Layzer's dismissal of them seems somewhat too peremptory.

Let me now turn to Layzer's discussion of the work on heritability of IQ. He assumes that most workers in behavior genetics are unaware of the difficulty of "disentangling the genotypic and environmental contributions to phenotypic variances" in IQ (*1*, p. 1263). A perusal of the literature, however, should convince anyone that many researchers are and have been fully aware of this difficulty (*2*). Layzer states (*1*, p. 1263) that "in adult subpopulations, IQ and environment are well known to be more or less strongly correlated." I am not sure what data he has in mind. Perhaps he is referring to the well-established relation between IQ and socioeconomic status. In respect to these data, although the magnitude of this correlation is usually high in any parental (that is, adult) population, it is not nearly so large in the filial generation (*3*). This shift may well mean that the problem is not as serious as Layzer thinks. Indeed, the analyses by Jinks and Fulker (*4*) and Jinks and Eaves (*5*) suggest that the problem of genotype-environment covariances may not be of such gravity as to render suspect all estimates of heritability. I grant the legitimacy of Layzer's argument that genotypes, regardless of the macroenvironments to which we may allocate them, may still be disposed to select microenvironments within these. I do not agree, however, that this means automatic defeat for the behavior geneticist. Indeed, the pioneering work of Harris (*6*) and of Wecker (*7*) on habitat choice in *Peromyscus* subspecies, to give one example, has been aimed precisely at this point.

Apart from the foregoing argument, Layzer also suggests that in any case the critical studies on human monozygotic twins reared apart (MZA) have all involved some selective placement such as to generate genotype-environment interactions. This is certainly true in some cases, but, in the particular study he chooses to cite—that by Burt (*8*)—the empirical correlation between socioeconomic status of one twin and socioeconomic status of the other twin is no different from zero. Jensen has computed such a correlation (.03), and I have verified it using a slightly different method. The fact that in most cases one member of each pair was reared by his natural parent is quite irrelevant to any conclusion. Consequently, I do not conclude, as Layzer does, that all the MZA data can be dismissed. Much the same applies to the fostering studies.

Layzer appears to favor the kind of design represented by the various intervention programs—for example, the Milwaukee Project (*1*, p. 1264). I would not argue against this. However, apart from the fact that most such projects (including especially the Milwaukee Project) have not produced lasting gains, they are still by no means free of the problems which Layzer sees as being inherent in the twin and fostering studies. They do not involve random allocation of genotypes to treatments, nor do they, to any degree, control for selection by genotypes of microenvironments. Certainly, on humane grounds, these kinds of enterprise must be considered desirable and essential. But it hardly seems likely that they will furnish us with answers of a precision and quality that are scientifically impeccable.

In conclusion, I must emphasize that Layzer's intelligent interest in the heritability of IQ should be welcomed by behavior geneticists. I welcome it myself. But I do consider that his assessment of more than 70 years' work on the problem, carried out by a great number of investigators using a variety of methods, is unduly harsh and overestimates the ambiguities that exist in the data.

WILLIAM R. THOMPSON
Department of Psychology,
Queen's University,
Kingston, Ontario, Canada K7L 3N6

References

1. D. Layzer, *Science* **183**, 1259 (1974).
2. See, for example, L. Ehrman, G. S. Omenn, E. Caspari, Eds., *Genetics, Environment, and Behavior* (Academic Press, New York, 1972); A. R. Jensen, *Educability and Group Differences* (Methuen, London, 1973).
3. C. Burt, *Br. J. Stat. Psychol.* **14**, 1 (19 S. Scarr-Salapatek, *Science* **174**, 1285 (1971); J. H. Waller, *Soc. Biol.* **18**, 252 (1971).

J. L. Jinks and D. W. Fulker, *Psychol. Bull.* **73**, 311 (1970).

5. J. L. Jinks and L. J. Eaves, *Nature (Lond.)* **248**, 287 (1974).
6. B. T. Harris, *Contrib. Lab. Vertebr. Biol. Univ. Mich.* **56**, 1953 (1952).

S. C. Wecker, *Ecol. Monogr.* **33**, 307 (1963).

8. C. Burt, *Br. J. Psychol.* **57**, 137 (1966).

9 April 1974; revised 29 July 1974

It is unfortunate that the phrase "Science or numerology?" in the title of Layzer's article (*1*) implies such scorn, for his own rigorous and serious attention to heritability is adequate proof that h^2 is not to be casually laughed off. The standard analyses used by geneticists and psychologists are obviously not those of a quack science, although there surely are, as Layzer asserts, some hidden assumptions that must be considered. And indeed Arthur Jensen, among others, has already taken careful note of certain of these (*2*).

There also appear to be some hidden assumptions implied by Layzer's own analysis. His first words pose the question: "To what extent can the development of basic cognitive skills be influenced by various kinds of environmental intervention?" (*2*, p. 1259). He then argues that h^2 is central to this question, and he proceeds to a mathematical analysis of h^2, showing that it is often difficult to tell, from such *broad* heritability, the amount of *narrow* heritability. He draws particular attention to two problems in such estimates, one caused by interaction and one by covariance. And he ends with praise for a particular environmental intervention, the so-called Milwaukee Project.

Taken as a whole, Layzer's article implies the following: that his mathematical criticisms of h^2 strengthen the environmentalist case against the hereditarian arguments of Jensen, Herrnstein (*3*), and others. His article, then, implies that any doubt cast upon heritability estimates can somehow bolster the sagging faith in such educational and social interventions. His own analysis, however, seen in statistical and logical terms, carries a strong refutation of any such optimism. It is the aim of this letter to make such paradoxes explicit.

Layzer's first major criticism of h^2, when applied to intellectual tests on human populations, is that there may be a substantial amount of "interaction" of genetic and environmental influences concealed in the usual large estimates. In formal terms, this possibility is represented (*1*, p. 1260) by his equation 5,

$$Var(P) = Var(G) + Var(E) + 2\,Cov(G,E) + I$$

where P stands for the measured, or phenotypic, trait, G is the contribution of direct genetic influences, E is the similar contribution of environment, $Cov(G,E)$ is the contribution of the correlation between heredity and environment, and I represents the interaction of heredity and environment. Layzer's valid point is that when estimates of heritability of IQ (or any other measure) are made in the usual way, h^2 may be inflated by such covariance and by such interaction.

For simplicity, let us consider these possibilities separately, beginning with the possible influences of interaction, or I in his equation. To understand the paradox presented by such interaction, let us examine the hypothetical data in Fig. 1. This figure is designed to show only the direct effects of G and E. In these fictitious data, genetic influences have added 20 IQ points on the high side and have subtracted 20 on the low side (seen in the marginal means at the bottom of the figure). In the same way, environmental influences have added (or subtracted) 10 IQ points in the rows. The cell means show that they are calculated only from the direct effects of the rows and the columns. For instance, the upper-right (high-high) cell has added 20 for G and 10 for E, for its mean of 130 IQ.

Now, by definition, any interaction of G and E will not change the marginal means. In order to keep the same G and E direct effects, we must keep the same row and column means. Suppose we add 5 IQ points, by some educational intervention, to the lower-left cell. Then to keep the figures in balance we must subtract 5 IQ points from the lower-right cell. Put another way, simply providing a "good" environment for all the low-environment subjects will produce an overall benefit only to the degree that the *direct* environmental effects are operative.

The first paradox, then, is this: No matter whether I is large or small in Layzer's equation, there should be no inference from his analysis of any support for the usual social and educational interventions. Indeed, such interventions are commonly based on some principle of "equalization," of making lower-class homes or schools more like upper-class homes or schools. In other words, the hopes of such programs are usually pinned on capturing the benefits of environment as found in current class differences. In such a context, the huge h^2 usually found, and the relatively small e^2, are meaningful indeed. And the statistical interaction of G and E offers no comfort, for any influence of such interaction is, by definition of I, as bad as it is good.

E (environmental effects) \ G (genetic effects)	Low	High	
High	90	130	110
Low	70	110	90
	80	120	100

Fig. 1. Illustrative table of genetic and environmental effects, where all effects are additive and there is no interaction of G and E. The marginal values are means.

The second major paradox of Layzer's analysis has to do with the importance he gives to $Cov(G,E)$ in his equation above. Indeed, Layzer has laid a logical trap for himself in the matter of covariance, and his argument exposes a very large inconsistency. Again, for simplicity, the I will now be ignored and the equation simplified.

We usually assume that "good environment" has something to do with socioeconomic status (SES), with cultural influences in the home and school, and with similar measurable signs of well-being. On the other hand, "good heredity" would mean genetic, innate endowments making for a desirable phenotype. In the case of IQ, such "good" environment and such "good" genotype would each contribute to higher measured IQ. By definition, $Cov(G,E)$ is dependent on the correlation between environment and genotype. Put differently, if such GE covariance holds, then the higher SES groups are, already, innately smarter than the lower SES groups. And the higher the GE covariance the greater is the genetic gulf between social classes.

But this is exactly what Jensen, Herrnstein, and other hereditarians argue: that there is, in fact, a genetic difference between SES levels, favoring the upper classes. And this is exactly what Layzer set out to refute, in his critical partitioning of the variance of the h^2 estimates. Considering only the covariance, then, he is led into a startling dilemma. Either h^2 has a substantial component of $Cov(G,E)$, or it does not. If it does not, then h^2 is to that extent a measure of direct genetic influences in IQ. If h^2 does include a substantial component of $Cov(G,E)$, then to that extent we grant the hereditarian thesis; that is, we acknowledge that social class is already partly determined by IQ genotype. Yet this is a flat contradiction to Layzer, who asserts (*1*, p. 1265):

> As long as systematic [cultural and environmental] differences remain and their effects cannot be reliably estimated, no valid inference can be drawn concerning genetic differences among races. Precisely the same arguments and conclusions apply to the interpretation of IQ differences between socioeconomic groups.

Thus we see that Layzer's analysis constitutes reductio ad absurdum, and carries its own refutation.

The 1950's and 1960's witnessed a boundless optimism about the potential benefits of a series of remedial programs aimed at lower-SES populations, and hundreds of millions were poured into experimental programs. A long string of disappointments has eroded our confidence, perhaps capped by the most rigorous large-scale experiment in educational history, conducted by the Office of Economic Opportunity (*4*). This test was of "performance contracting," but because of its diversity, generality, and unprecedented rigor of double-blind testing it served as a test, as well, of many behaviorist principles. And its failure was a major blow to such optimism (*5*). In fact, by 1970 it was becoming commonplace that remedial programs for the disadvantaged looked effective only so long as they were not subject to close scrutiny by critical outsiders.

Layzer continues this tradition by praising "the remarkable achievements of the Milwaukee Project" (*1*, p. 1264), which despite its being "now in its sixth year" (p. 1265) still had only the status of an "unpublished research report." In fact, the unpublished report by Heber (*6*) was obtained 2 years ago, after considerable difficulty, and was analyzed in the technical literature (*7*). Since that time, a 1972 "progress report" (*8*) has had some limited circulation, but it has not explained sufficiently the defects in design and reporting, nor given any theory which would account for the results, so different from those of other investigators. In brief the widely publicized gain of 30 IQ points has no clear scientific or practical meaning, since no educator can know how to replicate the effects. That Layzer's principal environmentalist evidence should have such status is, once again, an implicit refutation of his position (*9*).

ELLIS B. PAGE
Department of Educational Psychology, University of Connecticut, Storrs 06268

References and Notes

1. D. Layzer, *Science* **183**, 1259 (1974).
2. A. R. Jensen, *Educability and Group Differences* (Harper and Row, New York, 1974), especially pp. 366–375.
3. R. J. Herrnstein, *IQ in the Meritocracy* (Little, Brown, Boston, 1973).
4. *An Experiment in Performance Contracting: Summary of Preliminary Results*, OEO Pamphlet 3400-5 (Office of Economic Opportunity, Washington, D.C., 1972).
5. E. B. Page, *Educ. Psychol.* **9** (No. 3), 40 (1972); *Phi Delta Kappan* **54** (No. 2), 115 (1972).
6. R. Heber and H. Garber, "Rehabilitation of families at risk for mental retardation" (progress report, University of Wisconsin, Madison, October 1971).
7. E. B. Page, *Educ. Res.* **1** (No. 10), 8 (19 ; *ibid.* **2** (No. 4), 2 (1973).
8. R. Heber, H. Garber, S. Harrington, C. Hoffman, "Rehabilitation of families at risk for mental retardation" (progress report, University of Wisconsin, Madison, December 1972).
9. I am indebted to the referee of *Science* for helpful comments. This referee pointed out a third possibility in the paradoxes discussed: If there is a high broad heritability and low narrow heritability, and the cause is dominance and epistasis, then the situation is not responsive to selection. But here, as in the case of *GE* interaction, there is no support for environmental intervention.

9 April 1974; revised 8 July 1974

Layzer's generally excellent article on IQ scores and inheritance (*1*) obscures two fundamental statistical facts in the discussion at the start of page 1261 and reference 10.

The fundamental facts relate to random variables x,z having a joint distribution. (I leave aside here irrelevant qualifications of existence and measure theory, and follow Layzer in denoting expectation by $\mathcal{E}$.)

First, the predictor of z from x that minimizes mean square error of prediction is $\mathcal{E}(z|x)$, the conditional expectation of z given x. This familiar fact holds for each value of x and is a standard way of describing the notion of expectation.

Second, there is zero covariance between the minimum mean square error predictor and its residual, that is, between $\mathcal{E}(z|x)$ and $z - \mathcal{E}(z|x)$. A proof is a one-line application of the basic repeated expectations relation: $\mathcal{E}[\mathcal{E}(u_1|u_2)] = \mathcal{E}u_1$. These manipulations arise statistically in, for example, the so-called Rao-Blackwell theorem.

In Layzer's article, take z as $P(x,y)$, so that $G(x) = \mathcal{E}(P(x,y)|x)$ is $\mathcal{E}(z|x)$. The covariance between G and Layzer's R when x,y are independent is the covariance between G and P-G, and one may immediately apply the second basic fact above to see that the covariance is zero. There is no need for series expansions or similar heavy machinery.

WILLIAM KRUSKAL

Department of Statistics, University of Chicago, Chicago, Illinois 60637

References

1. D. Layzer, *Science* **183**, 1259 (1974).

18 April 1974

I will attempt to deal with Thompson's concise and orderly critique paragraph by paragraph.

Paragraph 3: The answer to the question posed here, "whether h^2 at y_1 should be greater or smaller than at y_2," is "neither"; $h^2 = 1$ in both cases, and the question would be equally trivial for any specified environmental ranges. I do not understand why Thompson regards this question as presenting an "interesting and potentially soluble problem." The point I actually made in this connection was the obvious but important one that the heritability of a trait for a given population and a given set of environmental conditions tells us nothing about the relative importance of genetic and environmental variations under a different set of environmental conditions.

Paragraph 4: Toward certain problems—for example, the problem of squaring the circle or the problem of constructing a perpetual motion machine—an "attitude of hopelessness" may be appropriate. In the natural sciences existing data and existing theories are inadequate to answer many, if not most, of the most pressing questions. Recognition of specific kinds of inadequacy is what usually initiates fruitful attempts to develop new theories and devise new experiments. If my technical arguments concerning the heritability of phenotypically plastic traits in natural human populations are sound, attempts to extract meaningful conclusions from data on IQ correlations are bound to fail, whatever the attitude of the data analyst. Optimism uninformed by technical insight is of little use to a scientist.

Paragraph 5: Systematic errors are, roughly speaking, errors that do not average out, errors introduced by a bias or disturbance of which the experimenter is unaware and that affect the data in a nonrandom way. Systematic errors are usually revealed by discordances between different methods of measuring the same quantity (for example, distance measurements by radar and by triangulation). When only one method of measuring a quantity exists, measurements made by this method must be assumed to contain systematic errors. This is always the case for "measurements" that, like IQ, are defined in a purely instrumental way. Such "measurements" may have diagnostic or predictive value but have no quantitative significance, as was recognized by S. S. Stevens. Seventy years of research into "the fundamental problems of reliability and validity" have yet to produce an instrumental "measurement" free from uncontrollable systematic errors of unknown magnitude, and I know of no reason to suppose that the next 70 years will be more productive in this respect. What is at issue is not the reliability and validity of psychometric procedures but the domain of applicability of a *biological* theory. Theories are not omnivorous; they do not find all kinds of data equally digestible, and their dietary constraints are, unhappily, rather rigid.

Paragraph 6: On the reliability of IQ tests I would call Thompson's attention to the recent study by McCall *et al.* cited in my article and in my reply to Page (below).

Paragraph 7: My intention was not merely to point out a difficulty that, as Thompson rightly remarks, has been discussed by many previous writers, but to argue that that difficulty—disentangling the genotypic and environmental contributions to the phenotypic variances of phenotypically plastic traits in natural human populations—is insoluble. I also emphasized that, for nonhuman populations, the problem could be solved in principle if environments could be sufficiently randomized with respect to genotypes; so I agree with Thompson's final remarks in this paragraph. I cannot agree, however, with his assessment of the importance of genotype-environment correlation in natural human populations. Recent work on human development in infancy and childhood assigns ever-greater importance, as regards cognitive development, to the interaction between mother and child during the earliest months and years. Since the mother also provides half of her child's genes, a substantial degree of correlation between genotype and aspects of the environment most relevant to cognitive development seems unavoidable.

Paragraph 8: In view of the discussion by Kamin [(*1*); see also Jensen (*2*)], Burt's twin data can no longer be regarded as admissible scientific evidence. Even if this were not the case Thompson's argument would be invalid, for the occupational status of the father (the characteristic by which "socioeconomic status" is represented in Burt's study) is surely not an environmental factor that strongly affects cognitive development.

Paragraph 9: Intervention studies do not seek to answer the same questions as heritability studies, hence they do not encounter the same methodological problems. Intervention studies deliberately alter the child's environment. The most successful of them also seek to improve the mother's environment and the quality of mother-child interactions. Questions of randomization and equalization present themselves only in connection with the formation of control groups, and in this context they present no insuperable problems. As to the effectiveness of intervention programs, it is literally true that most of them have not produced sustained cognitive gains. This generalization, however, emphatically does not apply to a particular class of intervention programs: those that "place major emphasis on involving the parent *directly* in activities fostering the child's development" (*3*). As Bronfenbrenner (*3*) has shown, the successful intervention studies are beginning to define a clear and consistent pattern. They are beginning to teach us why disadvantaged children do not realize their cognitive potentials and how their chances of doing so can be improved. This is science without numerology.

Page's remarks fall under two main heads. (i) He considers my critique of heritability analyses to be internally inconsistent and self-refuting, and asserts that, my arguments to the contrary notwithstanding, valid inferences concerning genetic differences between socioeconomic and racial groups *can* be drawn from published heritability analyses. (ii) He argues that the outcomes of "remedial programs aimed at lower-SES populations" have failed to produce durable results that stand up to "close scrutiny by critical outsiders."

I think Page is mistaken on both counts. Consider first his strictures on "remedial programs." These recall Arthur Jensen's famous dictum, "Compensatory education has been tried and it apparently has failed" (*4*). Such judgments are doubly flawed. In the first place, they are, so to speak, ungrammatical. "Education," rightly understood, cannot take a verb in the perfect tense, for it denotes an imperfect process in both senses of the word: education is never completed, and it always admits of improvement. The second flaw in the Page-Jensen obituary for compensatory education is that it is premature. Although most compensatory and intervention programs have indeed failed to produce durable results, a few have produced substantial and sustained gains in cognitive performance. Urie Bronfenbrenner, in a report to the Department of Health, Education, and Welfare (*3*), has pointed out that these successful programs have certain key features in common, features that the unsuccessful programs lack. Although Bronfenbrenner's tentative conclusions need to be confirmed by additional work, they afford rational grounds for believing that appropriate large-scale social and educational programs could wipe out functional illiteracy and innumeracy as effectively as public health programs have wiped out smallpox and diphtheria.

What no intervention study or compensatory program has so far produced is a simple and inexpensive remedy for cognitive deficits among the disadvantaged. Bronfenbrenner argues that an effective intervention program must be "ecological" in scope: it must "provide adequate health care, nutrition, housing, employment and opportunity and status for parenthood." Given such intervention, the available evidence indicates that "even children from severely deprived backgrounds of mothers with IQs below 70 or 80 are not doomed to inferiority by unalterable constraints either of heredity or environment." Bronfenbrenner adds, however, that "ecological intervention will require major changes in the institutions of our society."

Page mentions two particular studies: "performance contracting," which he describes as "the most rigorous large-scale experiment in educational history"; and Heber's Milwaukee Project, which he dismisses as having "no clear scientific or practical meaning." The failure of performance contracting, he says, was a "major blow" to "optimism about the potential benefits of remedial programs aimed at lower-SES populations." Its actual impact was considerably milder than this language might suggest. It did, perhaps, dampen the hope that operant conditioning in the classroom could overcome the effects of severe and sustained physical, emotional, and cognitive deprivation. (I say "perhaps" because the rigor of the experimental design did not, unfortunately, extend to its implementation.) But this hope does, after all, reflect a view of cognitive development that modern studies (*5*) had rendered exceedingly improbable long before performance contracting was sold to the Office of Economic Opportunity.

As to the Milwaukee Project, I have to admit that a careful reading of Page's published critiques (*6*) and of the lucid, detailed, comprehensive, and copiously documented report by Heber, Garber, Harrington, and Hoffman (*7*) has left me in a state of mystification concerning the substantive basis for his criticisms. In any event, Page errs in asserting that Heber's findings lack a theoretical framework and contradict the findings of other investigators. In the report cited earlier, Bronfenbrenner (*3*) makes precisely the opposite points. "Given our frame of reference," he writes, "the success [of Heber's program] is not unexpected since the program fulfills major requirements we have stipulated as essential or desirable for fostering the cognitive development of the young child." Bronfenbrenner cites a number of other recent studies, as well as a few older studies, that strongly support his hypotheses concerning cognitive development and are entirely consistent with Heber's results. Moreover, a recent longitudinal study of normal, home-reared, middle-class children found that "the average individual's range of IQ between $2\frac{1}{2}$ and 17 years of age was 28.5 IQ points, one of every three children displayed a progressive change of more than 30 points, and one in seven shifted more than 40 points" (*8*). In light of this finding, is it really so surprising that an intervention program as intensive, comprehensive, and meticulously well planned as Heber's should have produced (in some ways its least impressive result) an average IQ gain of 30 IQ points among the children of Black mothers with IQ's of 75 or less living in a severely depressed area of Milwaukee?

I come now to Page's comments on my critique of heritability analyses. I do not reproach Page for failing to grasp my mathematical arguments; perhaps they were not as clearly expressed as they might have been. But I am dismayed by his apparent failure to grasp the qualitative meaning of genotype-environment interaction in the context of human development and by his failure to understand that no valid inference concerning the genetic basis of differences in cognitive performance between social or ethnic groups can be drawn from heritability estimates.

Genotype-environment interaction is important whenever (i) a given environmental change produces substantially different phenotypic responses in individuals of different genotypes or different environmental histories, or (ii) a given genetic difference (as, for example, between fraternal twins) would have substantially different phenotypic consequences in substantially different environments. On biological grounds it is reasonable to assume that genotype-environment interaction is largely responsible for the variance of phenotypically plastic traits in natural populations. This view assigns equally strong roles to genetic and environmental variations; hence it is unpopular with both hereditarians and environmentalists. Moreover, it denies the possibility of separating the environmental and genetic components of the phenotypic variance by statistical analysis; hence it is unpopular with quantitative geneticists. Nevertheless, even when genotype-environment interaction is important for a given trait, it may be possible to derive meaningful heritability estimates for that trait. The main result enunciated in my article was that such estimates are possible in principle if and only if genotype and environment are uncorrelated. Since this condition is never met by phenotypically plastic traits in natural human populations, I concluded that meaningful heritability estimates cannot now be obtained for such traits.

Page argues that assuming genotype and environment to be correlated is tantamount to granting the hereditarian thesis that "the higher SES groups are, already, innately smarter than the lower SES groups." To pinpoint the fallacy in this argument, consider a phenotypically plastic trait that is easier to define and measure than intelligence: proficiency in the game of squash. Few people will deny that this proficiency is correlated with genetic factors (for example, genes specifying a predilection for strenuous forms of exercise). It is also undeniable that the general level of proficiency at squash is substantially greater among students and graduates of Ivy League colleges than among students and graduates of the Big Ten. Page and the authors whose views he cites with approval would, I hope, reject a genetic explanation for this systematic difference. Why, then, do they persist in interpreting systematic behavioral differences between social, economic, and racial groups as evidence for systematic genetic differences?

I thank Kruskal for his clarification.

DAVID LAYZER

Department of Astronomy,
Harvard University,
Cambridge, Massachusetts 02138

References

1. L. J. Kamin, *The Science and Politics of I.Q.* (Wiley, New York, 1974).
2. A. R. Jensen, *Behav. Genet.* **4**, 24 (1974).
3. U. Bronfenbrenner, "Is early intervention effective?" (unpublished condensed version of a report of the same name, Office of Child Development, Department of Health, Education, and Welfare, Washington, D.C., in press).
4. A. R. Jensen, *Harv. Educ. Rev.* **39**, 1 (1969).
5. See, for example, T. G. R. Bower, *Development in Infancy* (Freeman, San Francisco, 1974).
6. E. B. Page, *Educ. Res.* **1** (No. 10), 8 (19 *ibid.* **2** (No. 4), 2 (1973).
7. R. Heber, H. Garber, S. Harrington, C. Hoffman, "Rehabilitation of families at risk for mental retardation," (progress report, University of Wisconsin, December 1972).
8. R. B. McCall, M. I. Appelbaum, P. S. Hogarty, *Monogr. Soc. Res. Child Dev.* **38**, 1 (1973).

21 March 1975

The Lower "Petrologic Geotherm": A Transitory State

In a recent conceptual advance, MacGregor and Basu (*1*) have presented a "petrologic model" of the geotherm in the upper 200 km of the earth. This work is important for at least three reasons: (i) it is based on concepts different from those previously used in modeling thermal structure in the earth and so provides an independent check on these concepts; (ii) it is in general agreement with the earlier models for the upper 140 km (although with interesting changes of detail), thus deepening our understanding of this region; and (iii) it reveals a new feature (a steepening of the geotherm) below 140 km beneath continents, thus inaugurating a new discussion of the thermal structure in this region. Although MacGregor and Basu are aware that this feature cannot reflect the steady state, and although two transient mechanisms (*2, 3*) are mentioned, the experimental data are interpreted as a petrologic *model* of the geotherm, applicable to a typical subcontinental tectonic setting, and generalizable (with suitable evolutionary modification) to other such areas. My purpose in this technical comment is to argue, on very simple grounds, that the "geotherm" represented by the lower part of the petrologic model must represent an extremely unusual state of the asthenosphere and cannot represent any steady evolutionary development applicable to other times or places, especially to the "typical" state of the upper mantle.

The conservation of energy, in the case of heat flow in one dimension, is expressed as

$$\rho\, C_P \frac{dT}{dt} = \frac{dF}{dz} + Q \qquad (1)$$

where ρ is the density, C_P is the specific heat, T is the temperature at depth z and time t, F is the vertical heat flow, and Q is any heat source density. Equation 1 is valid for all materials (for example, inhomogeneous plastics) if F is suitably defined (below). The one-dimensional case is sufficient for this problem, as the indicated (*1*) horizontal temperature gradients are less than the vertical gradients by more than an order of magnitude. The heat flow is given by

$$F(t,z) = k_s \frac{dT}{dz} + F_{conv}$$

where the conductive term (the first term on the right) depends upon the thermal gradient and k_s, the thermal conductivity in the stationary state. The convective contribution F_{conv} is indicated only symbolically; it is always positive or zero. Hence

$$F(t,z) \geqslant k_s \frac{dT}{dz} \qquad (2)$$

in both liquids and solids. It has been shown (*4*) that k_s increases slowly in the upper 400 km, being always greater than 3×10^5 centimeter-gram-second units.

One can apply these equations in an elementary way to a temperature distribution showing upward curvature, as in the "petrologic model" (*1*). Neglecting possible heat sources Q for the moment, Eq. 2 substituted in Eq. 1 yields

$$\frac{dT}{dt} > \frac{k_s}{\bar{\rho}\, C_P} \frac{\Delta(dT/dz)}{\Delta z} \qquad (3)$$

Using a change of gradient $\Delta(dT/dz)$ of 16°C per kilometer over a depth interval Δz of 30 km [suggested by the data (*1*)] in Eq. 3 yields a minimum value $dT/dt > 10^{-5}$ °C per year. This indicates that temperature excesses (over the extrapolated lithospheric geotherm) of the order of 100°C would decay away (by heating of the lithosphere) in a maximum of (10 to 20) $\times 10^6$ years and possibly sooner; that is, if the petrologic model geotherm does represent true paleotemperatures just prior to surface emplacement (some 100×10^6 years ago) of the corresponding ultramafic rocks, that thermal structure has long since smoothed itself out. Conversely, it could not have existed for more than a few million years prior to the emplacement event without conductively heating up the lithosphere and removing the inflection. Hence it must be considered an extraordinary situation, not part of a steady evolutionary development, and not generalizable to other areas in similar tectonic settings (for example, similar distances from spreading centers). In effect, it constitutes petrographic evidence of a transient (or mobile) anomalously hot spot in the mantle rather than a representative geotherm.

The neglected source term Q does not affect this conclusion. The time scale for the conduction of heat into the lithosphere does not depend on the source of the heat, be it convectively transported from below or internally generated by radioactivity or viscous dissipation. Only a negative heat source (a heat sink) at the top of the asthenosphere could maintain a concave-upward geotherm for significant times. The only heat sinks available are endothermic chemical reactions, such as melting or dehydration, and descending diapirs. Considering first the endothermic reactions, the reaction rate required to maintain an inflection in the geotherm is easily calculable. It is more instructive, however, to estimate the steady-state rate of accumulation of reaction products, since this quantity is independent of Δz, the interval of upward curvature of $T(z)$. This production rate is easily shown to be

$$p > \frac{k_s}{L} \Delta(dT/dz) \qquad (4)$$

where L is the latent heat of the reaction. Using $L = 100$ cal/g (for the melting of forsterite), one calculates $p > 0.3$ g/year per square centimeter of horizontal area, corresponding to a column of reaction