Fallacious Use of Regression Effects in the I.Q. Controversy*

Brian Mackenzie, University of Tasmania

ABSTRACT

Regression to the population mean in I.Q. scores has been taken by H.J. Eysenck, A.R. Jensen, and other writers to provide evidence for genetic determination of individual and racial I.Q. differences. However, regression is a purely statistical phenomenon, and as such provides no evidence for either genetic or environmental determination of I.Q. The statistical basis of regression is explained, the misuse of regression effects in the literature is reviewed, and arguments based on such misuse are criticized. It is suggested that the fallacious use of regression effects is partly a result of the isolation of biometrical from Mendelian genetics.

In 1886, Francis Galton announced to the world his Law of Filial Regression: "When Mid-Parents are taller than mediocrity, their Children tend to be shorter than they. When Mid-Parents are shorter than mediocrity, their Children tend to be taller than they" (Galton, 1886, quoted in Pearson, 1930, p. 16). For Galton, this regression towards mediocrity was a biological process caused by blending of the germ plasm over generations. As Pearson put it, regression was for Galton "the result of the influence of parental heredity pulling the offspring . . . towards the parental value and the mediocrity of the more distant ancestry pulling towards its own value" (Pearson, 1930, p. 20). When Pearson came to work out the mathematics of correlation (Pearson, 1896), he gently corrected his mentor's mistake. As he put it later, regression, and the Law of Ancestral Inheritance based upon it, "is not a biological hypothesis, but the mathematical expression of statistical variates which obey . . . certain forms of frequency distribution, these being maintained in successive generations" (Pearson, 1930, p. 21). In short, regression is a statistical phenomenon, not a biological one.

Pearson's lesson has not been well learned. A substantial number of psychologists continue to regard regression as a genuine biological phenomenon. While they do not repeat Galton's error of basing regression on blending inheritance, they do reach other conclusions about a presumed genetic basis for regression that are equally erroneous. These errors play a significant role in the current controversy on genetic and environmental influences on I.Q. For this reason, it will be worthwhile to analyze them and, in so doing, try to prevent their recurrence.

* Requests for reprints should be addressed to Brian Mackenzie, Department of Psychology, University of Tasmania, G.P.O. Box 252C, Hobart, Tasmania, 7001.
THE SIMPLEST CASE: PARENT-OFFSPRING REGRESSION

The most explicit recent statement by a psychologist on the genetic causation and hence the biological reality of regression is one by Eysenck (1971). In discussing Cyril Burt's data on the relationship of I.Q. to occupational class, he notes that Burt's 'Higher professional' group had a mean I.Q. of 140, his 'Unskilled' group of 85. He then makes the following observation about the I.Q.s of their children:

It will be noticed that something very interesting has taken place. We now have regression to the mean; the children of our higher professional fathers have a mean I.Q. of only 121, that is, they have regressed half way to the mean of the whole population, which is of course 100. Similarly, the children of the unskilled parents have an I.Q. of 93; they have regressed upwards and roughly half way towards the population mean. This regression to the mean is a phenomenon well known in genetics, and characteristic of traits markedly influenced by genetic causes; environment would favour the children of the higher professional fathers, and disfavour those of unskilled working-class fathers, tending to make the difference between them even greater than that observed between their fathers. Clearly this is not what happens; regression presents strong evidence for genetic determination of I.Q differences (Eysenck, 1971, pp. 67-8).

The linkage of regression with genetic influences is clear and unambiguous; genetic determination is expected to produce regression, environmental determination is not.

In apparent contrast to Eysenck's position is a statement in a recent advanced textbook on human genetics (Vogel & Motulsky, 1979). In discussing Galton's early data on the stature (height) of parents and children, the authors write as follows:

With additive gene action, the children's mean is expected to be exactly one-half between parents' values, i.e., should be identical to the midparent value. This however, is not the case. Instead, the data show something else: if the midparent value is higher than the population mean, the children's mean is lower than that of their parents. On the other hand, if the midparent value is lower than the population mean, the children's mean is higher. Hence, as a rule, the children's mean tends to deviate from the parents' mean in the direction of the population mean.

This phenomenon was observed by Galton and named "regression to the mean." It can also be shown in other, similarly continuously distributed characters.

What is the reason for this divergence from genetic expectations? Individuals who can be ranged at the extremes of a distribution curve presumably obtain not only the genetic factors that make for the extreme phenotype but probably have benefited in addition from unusual environmental circumstances. Furthermore, specific gene-gene and gene-environmental interactions may have been operative in actualizing their extreme phenotypes. Their children on the average are less likely to have benefited from the special environmental influences and gene-environmental
interaction that placed the parent in the extreme categories. Their phenotypic values are, therefore, more likely to be similar to the mean of the population — a regression to the mean (Vogel & Motulsky, 1979, p. 154).

Vogel and Motulsky differ from Eysenck in the specific explanation they offer for regression. They maintain that the phenomenon of regression is contrary to what would be expected by a simple genetic model (i.e., with additive gene action) and that environmental and interactive influences must therefore be invoked to account for it. Thus the psychologist, finding an effect which cannot be accounted for on the basis of differential experience, proposes a largely genetic explanation; the geneticists, finding an effect which cannot be accounted for on the basis of established genetic models, propose a largely environmental explanation.

Vogel and Motulsky agree with Eysenck in the most important point however, that regression requires an explanation by either genetic or environmental factors, that is, that it calls for specific explanation in terms of the casual factors underlying growth and development. This view is fallacious. The explanation of why it is fallacious may be best offered now, before we examine more complicated variants on it.

WHY IS THERE REGRESSION TOWARDS THE MEAN?

The fact that Eysenck’s and Vogel and Motulsky’s interpretation of regression is incorrect can be shown easily enough. With regard to Eysenck, it is sufficient to point out that inter-generational regression works in both directions. If fathers with I.Q.s of 140 have children with I.Q.s that average only 120, so likewise, do children with I.Q.s of 140 have fathers with I.Q.s that average only 120. The phenomenon of inter-generational regression to the mean does not arise from any influences of parents on their children, whether genetic or environmental. It arises from the limited similarity of parents and children, and can be seen in both directions equally. With regard to Vogel and Motulsky, it is sufficient to point out that regression does not occur only, or even mainly, at the extremes of the distribution. The amount of regression is of course greatest at the extremes, but is a constant proportion of the predictor score’s deviation from the mean. Vogel and Motulsky report a midparent-child correlation for height, based on Galton’s data, of 0.449. If this value is accurate, then parents whose height is 10 inches above the mean will have children whose height averages 4.49 inches above the mean; parents whose height is 1 inch above the mean will have children whose height averages 0.449 inches above the mean.¹ Strict proportionality is preserved, and again,

¹ Assuming equal means and variances in each generation, and disregarding sampling error.
parents' heights can be predicted from knowledge of their children's as easily as the other way around, and with the same degree of regression occurring. This is not to deny that there may well be casual factors such as Vogel and Motulsky propose, that influence the height only of individuals at the extremes of the distribution. The relationship of the height of these individuals to the height both of their parents and of their children, however, will be other than what would be predicted by simple regression. That is, it will reduce the accuracy of prediction at the extremes of the distribution.

What, then, does 'cause' regression ('cause' being placed in quotes since the explanation does not involve causation in the usual sense)? The full and proper explanation of regression is contained in the mathematical theory of least squares. However, since some writers interpret the concept incorrectly, despite familiarity with the statistical procedures required for its use, a simplified conceptual explanation may be appropriate. In what follows, we will assume an X and a Y distribution, each with mean of 100 and standard deviation of 15 (such as might occur, for instance, with I.Q. of siblings), and an XY correlation of 0.4. Predictions will be made from the X scores; the predicted Y scores will be designated as Y'.

If we follow the usual and unproblematic statistical procedure to generate a Y' distribution, it will have three salient features: the rank order of the Y' scores will be the same as the rank order of the corresponding X scores, the means of the two distributions will be equal, and the standard deviation of the Y' distribution will be r_{xy} times the standard deviation of the X distribution. In our numerical example, the Y' distribution will have a standard deviation of 0.4 x 15, or 6.0. This reduced standard deviation corresponds to the proportion of the variance in Y that can be predicted from X. With a correlation of 0.4, only 16% of the variance can be predicted, and the variance of the Y' scores (6^2 = 36) is 16% of the variance of either the X or the actual Y scores (15^2 = 225).

What it means for the Y' scores to have a smaller variance than the X scores is that the Y' scores are less spread out, or more tightly bunched together, around their mean than the X scores are; variance is the measure of 'spread-outedness' or 'bunched-togetherness'. Since the Y' scores are clustered more tightly around the mean than the X scores are, while preserving the same rank order and mean, each one of

2. Equal means and variances are not necessary, but simplify the discussion, since they permit the effects to be expressed in raw or deviation scores, rather than in standard scores.

3. According to the general formula:

\( Y' = r_{xy} \left( \frac{S_Y}{S_X} \right) (x - \bar{x}) + \bar{Y} \)

Since, in the example, the means and standard deviations are the same for the X and Y distributions, the formula can be simplified verbally to: the correlation, multiplied by the X score's deviation from the mean, plus the mean.
them will be closer to that mean than the corresponding X score is. The amount by which each Y' score will be closer to the mean than the corresponding X score is the amount of its regression to the mean. This is all that regression to the mean does or possibly can signify: it is the shrinkage, or reduction in variance, of the Y' distribution compared to the X distribution, due to the inability of the X distribution to account for some part of the variance in Y. It does not involve or permit any separate casual factors that push extreme scores closer to the middle; as pointed out, any such casual factors would produce an effect separate from the general phenomenon of regression.

Fig. 1. Regression as the Reduction in Variance of a predicted (Y') Distribution

The relationship between the X and the Y' distributions is illustrated in Figure 1. For an X score at the mean of the X distribution, the corresponding Y' score (that is, in the example, the score predicted for the sibling) is at the mean of the Y' distribution. Likewise, for an X score lying two standard deviations above the mean of the X distribution, the corresponding Y' score lies two standard deviations above the mean of
the Y' distribution — but the standard deviations in the Y' distribution are smaller, so the score is closer to the mean. The X score is 30 points (2 x 15) above the mean, while the Y' score is just 12 points (2 x 6) above it.

The regression remains when we consider the actual Y scores, rather than the Y' scores, because the error variance is by definition uncorrelated with Y'. It thus increases the 'spread-outedness' of the Y scores to the level of their total variance, but does not systematically raise or lower any of them. Indeed, the necessity to add the error variance, that is, variance that is not accounted for by X, shows in another simple way why regression must occur. If there were no regression, so that the predicted or Y' scores were the same as the X scores, the Y' distribution would have the same variance as the X distribution and there would be no room left for the error variance. If the error variance were nevertheless added in, the complete Y distribution would have a greater variance than the X distribution, and since we started with X and Y distributions with the same variance, this result is unallowable.

The relationship between the Y' and the Y distributions is illustrated in Figure 2. All those Y' scores which lie two standard deviations (12

Fig. 2. Continued Presence of Regression in the Y distribution
points) above the mean of the Y' distribution generate, in the total Y
distribution, a distribution of their own. This smaller distribution,
which is one component of the total Y distribution, has a mean that is
12 points above the overall Y mean and a standard deviation that is the
square root of the error variance, that is, $\sqrt{225 - 36}$ or 13.75. The same
is true of the other component distributions, shown in light lines in
Figure 2. Their means are equal to the corresponding Y' scores, and
their standard deviations are all 13.75. When the component distribu-
tions are combined, they become the Y distribution. The error variance
(13.75) and the variance of the Y' distribution ($6^2$) sum to produce the
variance of the Y distribution ($15^2$). The respective standard deviations
of course do not add up in this way; variances alone are additive.

All this is elementary, but is necessitated by the insistence of the
writers quoted on seeking a specific causal explanation for a
phenomenon that is a formal consequence of the mathematics of pre-
diction. It is curious that Eysenck in particular should do so, since he
also refers to the presumably meaningless but "very high correlation
over the first seventy years of this century between the number of iron
ingots shipped from Pennsylvania to California and the number of
registered prostitutes in Buenos Aires (Eysenck, 1971, p. 80)."
Regression results from correlation, not independently from any causal
factors, and affects prediction of the number of prostitutes or iron
ingots in exactly the same way as prediction of I.Q. or height.

REGRESSION AND RACE DIFFERENCES

Most writers who assume the biological reality of regression do not do
so in so simple a way as Eysenck. The assumption is more often made
in a slightly subtler way, as providing the basis for an explanation of
race differences in intelligence. At least three prominent researchers
have claimed that regression effects can be used in this way (Scarr-

That there are some race differences in measured intelligence, par-
ticularly between blacks and whites in the U.S.A., is not presently con-
troversial. That is, there is no doubt that many investigators have
found that blacks in the U.S.A. achieve lower average scores on a
variety of I.Q. tests than whites do. The dispute is over whether this
mean difference in I.Q. is due entirely to environmental factors, as
claimed by Mercer and Brown (1973), or has a substantial genetic com-
ponent, as claimed by Jensen (e.g., 1973a, 1973b) and others.

The relevance of regression effects to this controversy is most
clearly and baldly stated by Scarr-Salapatek:

Regression effects can be predicted to differ for blacks and whites if the two
races indeed have genetically different population means. If the population
mean for blacks is 15 IQ points lower than that of whites, then the offspring
of high IQ black parents should show greater regression (toward a lower population mean) than the offspring of whites of equally high IQ. Similarly, the offspring of low-IQ black parents should show less regression than those of white parents of equally low IQ (Scarr-Salapatek, 1971a, p. 1226). However, the most extensive appeal to regression effects as evidence for the hypothesis of interracial genetic differences in I.Q. is made by Jensen (1972, 1973a, 1973b). As he makes use of such effects in several different contexts to justify his hypothesis, his use of them will be documented fairly fully, even though he makes essentially the same case each time.

In his famous 1969 paper, ‘How much can we boost IQ and scholastic achievement?’ (reprinted in Jensen, 1972), Jensen discusses the findings of Heber (1968) concerning incidence of mental retardation, defined as an I.Q. below 75, in U.S. blacks and whites. Black parents had a higher proportion of retarded offspring than white parents, at every level of socioeconomic status. However, the preponderance of black retardates increased with increased SES (although the absolute proportion decreased). At the lowest SES level, blacks had 5.5 times as many low I.Q. children as whites (per 100 children). At the highest SES level, blacks had 13.6 times as many as whites. Jensen interprets this finding as follows:

If environmental factors were mainly responsible for producing such differences, one should expect a lesser Negro-white discrepancy at the upper SES levels. A genetic hypothesis, on the other hand, would predict this effect, since the higher SES Negro offspring would be regressing to a lower population mean than their white counterparts in SES, and consequently a larger proportion of the lower tail of the distribution of genotypes for Negroes would fall below the value that generally results in phenotypic IQs below 75 (Jensen, 1972, p. 164).

On the next page, Jensen cites evidence for the general phenomenon of:

much greater 'regression-to-the-mean' (from parents to their children) for high status occupations in the case of Negroes than in the case of whites. None of these findings is at all surprising from the standpoint of a genetic hypothesis, of which an intrinsic feature is Galton's 'law of filial regression' (Jensen, 1972, p. 165).

In his Educational differences, Jensen (1973a) reanalyzes data from Scarr-Salapatek (1971b) to show that the children of high SES black parents have lower scores on an aptitude test battery than the children of low SES white parents ('high' and 'low' defined as above and below the median for the combined sample of black and white parents). Jensen comments:

This finding seems more difficult to reconcile with a strictly environmental explanation of the mean racial difference in test scores than with a genetic
interpretation which invokes the well-established phenomenon of regression toward the population mean (Jensen, 1973a, p. 416).

He then reproduces Scarr-Salapatek’s statement quoted above, and continues:

In other words, on the average, an offspring genetically is closer to its population mean than are its parents, and by a fairly precise amount. Accordingly, it would be predicted that upper status Negro children should, on the average, regress downward toward the Negro population mean IQ of about 85, while lower status white children would regress upward toward the white population mean of about 100. In the downward and upward regression, the two groups’ means could cross each other, the lower status whites thereby being slightly above the upper status Negroes (Jensen, 1973a, pp. 416-417).

Elsewhere in the same volume, Jensen (1973a, p. 395) introduces the concept of parent-offspring regression in I.Q. as a general phenomenon, and comments, “This phenomenon, discovered by Sir Francis Galton, is called ‘regression towards the mean’, and it holds true for height and other inherited physical traits as well as the I.Q.” In this statement, Jensen appears to follow Eysenck in affirning the biological reality of regression in general, but he is less explicit and does not say that it does not occur with environmentally determined traits.

In his Educability and group differences, Jensen (1973b, pp. 185-186) repeats his discussion of Scarr-Salapatek’s data, as quoted above, with only minor changes in wording. Later in the same volume, he proposes an extensive study, similar to Scarr-Salapatek’s but more detailed, in which the I.Q.s of both parents and children are tested in a sample of high SES blacks and low SES whites. His predictions from environmental and genetic models are:

If these SES factors are more important determinants of IQ than genetic factors, there can be no doubt that the predicted result should be a much higher mean IQ for the upper SES Negro children than for the lower SES white children. . . Since there is some regression toward the population mean from parent’s IQ to child’s IQ, a genetic theory of the racial intelligence difference would predict that Negro and white children should regress toward different population means. . . Because the Negro population mean is about one standard deviation below the white population mean, the mean IQs of our two hypothetical groups of children would be much closer together than if we compared the mean IQs of low and high SES white children, and this should be so, according to our genetic hypothesis, even if the high SES Negro and white parents were perfectly matched for IQ (Jensen, 1973b, p. 239).

The only limitation he places on the genetic prediction is that:

The conformity of actual data to the predictions from this genetic model will, of course, be attenuated to the degree that the parents’ and offsprings’
environments have been dissimilar with respect to factors influencing mental development (Jensen, 1973b, p. 239).

Even this single limitation is avoided in an example given earlier in the volume, which also contains Jensen’s most detailed elaboration of the presumed genetic significance of regression. He writes:

The correlation among siblings of close to 0.40 on the Lorge-Thorndike Intelligence Tests in both the white and the Negro samples has an interesting consequence which may seem puzzling from the standpoint of a strictly environmental theory. It is entirely expected if one assumes a genetic model of intragroup and intergroup differences. This is the phenomenon of sibling regression toward the population mean... Genetic theory predicts the precise amount of regression...

[If we match a number of Negro and white children for IQ and then look at the IQs of their full siblings with whom they were reared, we find... [that] the Negro siblings average some seven to ten points lower than the white siblings. Also, the higher we go on the IQ scale for selecting the Negro and white children to be matched, the greater is the absolute amount of regression shown by the IQs of the siblings. For example, if we match Negro and white children with IQs of 120, the Negro siblings will average close to 100, the white siblings close to 110. The siblings of both groups have regressed approximately halfway to their respective population means and not to the mean of the combined populations. The same thing is found, of course, if we match children from the lower end of the IQ scale. Negro and white children matched for, say, IQ 70 will have siblings whose average IQs are about 78 for the Negroes and 85 for the whites. In each case the amount of regression is consistent with the genetic prediction. The regression line, we find, shows no significant departure from linearity throughout the range from IQ 50 to 150. This very regular phenomenon seems difficult to reconcile with any strictly environmental theory of the causation of individual differences in IQ that has yet been proposed (Jensen, 1973b, pp. 117-118).

From this last and most detailed statement, in conjunction with the ones previously quoted, we can extract three components of the regression effect which Jensen cites as evidence for the substantial genetic determination of race differences in I.Q.4 In the order of the frequency with which he cites them, these are:

1. The regression is in each case to the “respective population means and not to the mean of the combined populations.” This is also the single point made by Scarr-Salapatek.
2. “The precise amount of regression” can be predicted by genetic theory.
3. The regression is consistently linear; that is, “The regression line... shows no significant departure from linearity throughout the range from IQ 50 to 150.”

4. As previously noted, one of Jensen’s statements may also be intended as using parent-offspring regression as evidence for genetic determination of I.Q. within racial groups. Jensen’s intentions are not so clear in this instance however, and in any case the matter was fully dealt with earlier in this paper.
In each case, the use of the regression effect as evidence for the genetic
determination of racial differences, or indeed as evidence for anything,
is fallacious. That it is fallacious may be apparent from the previous dis-
cussion of regression generally, but will in addition be shown briefly for
each claim.

With regard to the choice of population mean to which scores will
regress, it is clear that there is no choice involved: scores will regress to
that population mean from which their deviations are measured. The
empirically assessed correlation for each group unambiguously deter-
mines the amount of regression, and the empirically assessed mean for
each group is the mean to which regression will occur. As long as the
two groups have different means, it is inconceivable that scores from
the two groups should regress “to the mean of the combined popula-
tions.” But this unfailing tendency to regress to the correct mean can
be overcome. If we wanted the scores to regress to the combined mean
in Jensen’s example, we would need only measure the sibling correla-
tion in the combined sample rather than in each racial sample sepa-
rately. Then the scores would indeed regress to the combined mean,
presumably midway between 85 and 100. The amount of regression,
however, would be greater, because the sibling correlation in the com-
bined sample would be lower than in each sample considered sepa-
rately, due to the heterogeneity of the samples. The heterogeneity, in
turn, consists in this case simply in the different means for the two
groups considered separately; it is that difference which we start with
and which Jensen attempts to account for. But however closely one
inspects the purely statistical consequences of this difference in means,
one will not thereby arrive at an explanation for it or at evidence to
back up a preferred explanation. The observed difference in means,
which is completely responsible for the observed difference in regres-
sion, is a phenotypic difference, not a genotypic one. Whether there are
genotypic differences that account for it is indeed the question, but the
differential regression effect will persist as long as the phenotypic
difference does, and will disappear as soon as the phenotypic difference
does, whether there are differences in genotype or not (for further
elaboration see Mackenzie, 1980).

With regard to “the precise amount of regression”, it depends in
each population solely on the precise amount of correlation and, to
revert to an earlier example, is as precise for iron ingots and prostitutes
as it is for I.Q. and height. As regression is a formal consequence of
correlation, no theory can in principle predict it as precisely as the
measured correlation can. What, then, is the role of genetic theory?
Vogel and Motulsky (1979), despite their earlier misinterpretation of
regression (which does not recur in the rest of their work), state the
relationship clearly:

In quantitative traits, the similarities between relatives are usually
expressed and measured as correlation coefficients. Such analyses are
strictly empirical and require no genetic concepts. The conceptual difficulties begin if we try to interpret empirical risk figures or correlations between relatives in terms of genetic variability. Theoretically, certain correlation coefficients can be derived from assumptions regarding the degree of relationship between two persons and the degree of dominance of these genes. The theoretical expectations can then be compared with those empirically observed, and from this comparison, the "heritability", i.e., the fraction of variability contributed by genes, can be estimated (Vogel & Motulsky, 1979, p. 480).

That is, the correlation between family members (which is responsible for the regression) is part of the data on which genetic theory works. Rather than genetic theory predicting the amount of regression, that which is responsible for the regression dictates to genetic theory the amount of genetic influence.

Finally, there is the linearity of the regression line. By this, Jensen presumably means that throughout the range of I.Q. 50 to 150, a linear function fits the paired data better than a curvilinear function. While this is no doubt true of Jensen's data, it is not clear why he should regard the fact as evidence for genetic determination of I.Q. differences, whether within or between races. What genetic theory can account for is certain specific departures from linearity, which occur because the distribution of I.Q. in the population has higher frequencies at the high and low ends of the scale than would be expected from a perfect normal distribution. Some of the deviations from normality at the low end of the scale can be accounted for genetically by the action of specific chromosomal abnormalities, such as those responsible for Down's syndrome, phenylketonuria, and other causes of severe retardation (others are due to exogenous causes, such as infantile injuries or illnesses that affect the brain). These mechanisms, that is, account for why there are more individuals with very low I.Q.s than would be expected if I.Q. were distributed normally. In the well known survey by Roberts (1952), siblings of such severely retarded individuals were found to have a mean I.Q. of close to 100, while siblings of mildly retarded individuals had a mean I.Q. of close to 80. In terms of regression, Roberts' finding indicates that the I.Q.s of the mildly retarded sibships fit the regression line about as well as do the I.Q.s of sibships in the population at large, while the I.Q.s of the severely retarded sibships do not; i.e., they depart from linearity. Genetic and other mechanisms account for this departure from linearity by identifying the specific influences on I.Q. that operate only at the low end of the scale. They do not, however, account for the mass of scores that do not depart systematically from linearity. At most, the linearity of regression sug-

5. "Presumably", because Jensen actually refers to the linearity of the regression line (which, as a straight line, can hardly depart from linearity) rather than to the linearity of regression.
gests only that the determinants of I.Q., like I.Q. scores themselves, are continuously and normally distributed within a wide range. It does not throw any light on what those determinants might be.

CRITICISMS OF THE REGRESSION ARGUMENT, AND A REPLY

The positions of Eysenck, Scarr-Salapatek, and Jensen have received a small amount of critical notice. Thoday (1973), reviewing Jensen’s *Educability and group differences*, summarized Jensen’s use of the differential regression effect as evidence for genetic racial differences and commented:

For some time, I fell into the same trap. But it is a trap, for populations must regress to their own mean whatever the cause, genetic or environmental, of the mean differences between the populations. This evidence is therefore as compatible with explanation in terms of the environmental factor $X$ as in terms of the genetic hypothesis. It adds nothing whatsoever to the strength of the genetic hypothesis (Thoday, 1973, p. 419).

Thoday does not, however, explain why populations must regress to their own mean, perhaps assuming that the reason is obvious.

A more detailed discussion of regression is provided by Loehlin, Lindzey, and Spuhler (1975). They begin, like Vogel and Motulsky (1979), by attempting to give a causal explanation for parent-offspring regression in height, but their explanation is so cautious and hedged with qualifications as to make it unobjectionable. They emphasize that some genetic and some environmental influences are shared between fathers and sons, and that some of both kinds are not, and conclude:

The regression depends on the proportion of factors that are shared, that is, on the correlation between father and son, and not on whether the shared factors are genetic or environmental. A father-son correlation of .50 predicts the same regression whether it is derived from shared diet or shared genes. The purely statistical character of the regression phenomenon may be appreciated by noting that it operates just as strongly in predicting the height of fathers from that of their sons as it does in the other direction (Loehlin et al., 1975, p. 175).

Then they consider the use that previous writers have made of regression effects and, citing some of the examples quoted above, note:

Since regression depends on the observed correlation and not on the heritability, one might well ask why Jensen (1969, p. 84) and others (e.g., Scarr-Salapatek, 1971a, p. 1226; Eysenck, 1971, p. 64) express the view that evidences of parent-child IQ regression fit genetic hypotheses better than environmental ones. This would seem largely equivalent to a claim that genetic hypotheses better predict the size of observed parent-child IQ correlations than do environmental hypotheses (Loehlin et al., 1975, p. 176).

While the description of regression offered by Loehlin et al. is fully
accurate, their interpretation of other writers’ use of the concept is overly charitable and potentially misleading. Although the two claims that Loehlin et al. regard as equivalent are, in fact, equivalent, they are not treated as equivalent by Eysenck or by Jensen. Both those authors discuss regression quite independently of correlation, and use the existence of regression effects, rather than just the accurate prediction of their size, as evidence for genetic determination. Furthermore, Jensen and Scarr-Salapatek both interpret regression effects as evidence for the genetic determination of racial differences, and in this case there can be no question of the claim being simply equivalent to a claim about the prediction of within-family correlations.

Even if the two claims described by Loehlin et al. were treated as equivalent, they would of course still be fallacious. As we have seen, prediction of the size of the within-family correlation or the amount of regression has no role as evidence for genetic determination, since the correlation is part of the data from which heritability is estimated. However, the two fallacies (using the occurrence of regression as evidence, and using the precise amount of regression as evidence) should be clearly distinguished, since they can be employed independently. The failure of Loehlin et al. clearly to distinguish the two could give rise to the erroneous impression that regression effects can have some status as evidence after all, as long as we avoid concentrating on those that arise wholly within families. Such a position is taken, partly in reply to Loehlin et al., by Vernon, in his monograph, *Intelligence: Heredity and environment* (1979).

Vernon discusses regression effects in two separate contexts: genetic influences on individual differences and genetic influences on group differences, with the two discussions separated by over 50 pages. His

6. Cf. Eysenck’s very clear claim quoted above, that environmental determination would make children’s I.Q. scores deviate from their parents’ scores in the opposite direction to that which characterizes regression. Eysenck elaborates this point at some length (1971, pp. 68-69).

7. There are exceptions to this procedure, as Loehlin et al. (1975, p. 176) note, in which the environmental contribution to the correlation is estimated a priori, usually at zero, and the correlation is predicted solely on a genetic basis. The difference is only in procedure, however, and not in the logic of the case, since the predicted correlation still has to be compared with the obtained one for a heritability estimate to be made.

Other exceptions are possible. Heritability estimates could, in principle, be based on kinship correlations other than the ones of interest, which could then be predicted in advance of measurement of them. However, this procedure is not advocated by Jensen in support of his claim that the amount of regression can be predicted by genetic theory, and has not in fact been generally adopted in human genetics, for a very good reason: different kinship correlations sometimes produce different heritability estimates. Jencks (1972) computed separate heritability estimates for I.Q. on the basis of several distinct kinship correlations, and found that they varied widely, from a minimum of less than .25 to a maximum of .76. Loehlin et al. (1975) recalculated some of Jencks’ estimates, using alternative assumptions in the path analysis, and found that they could reduce but not eliminate the variability.
Fallacious Use of Regression Effects

The repeated misuse of an elementary statistical concept, by psychologists who are highly trained in statistics, is a matter of some concern. What can account for it? In an earlier paper (Mackenzie, 1980), the present author suggested that the misuse of regression and other technical concepts provided "an extreme and transparent example of the scientism and methodolatry that has sometimes characterized our field: if it looks scientific, it must be scientific, and if we can clothe our ideas in the trappings of scientific method, then those ideas must have the weight of scientific method behind them." To make this suggestion more precise, it should be emphasized that the methodological error occurs at an earlier stage than that of the justification of preferred conclusions. Specifically, it may be suggested that the confusion over regression is symptomatic of the continuing isolation of biometrical from Mendelian genetics. Fisher's (1918) Mendelian interpretation of correlated traits is often taken to have signalled either the conceptual victory of Mendelian genetics over Pearson's biometrical statistical approach, or else the amalgamation of the two; Fisher showed that the continuous traits studied by the biometricians could be interpreted as resulting from the additive action of a large number of Mendelian genes. Rather than bringing the two approaches closer together, however, Fisher's analysis seems to have been taken by some psychologists implicitly to guarantee the autonomy of biometrical analysis. It assures them that any relationship that can be described by biometrical techniques must be due to some kind of gene action. Since the techniques of biometrical analysis are largely continuous with the statistical procedures in which psychologists are trained (and indeed, the statistical procedures were in many cases developed for biometrical purposes), it may seem natural to use them directly to demonstrate genetic influence. The misinterpretation of regression stems, on this account, from concentrating on it as a biometrical measurement rather than as a formal consequence of a mathematical model. With the former emphasis, it is natural to assume that regression must have a genetic cause; is that not what Fisher showed about biometrically described traits in general? The answer, of course, is, no, it is not; but unless the
formal cause of regression is pointed out in detail, the assumption may be an easy one to make.

The isolation of biometrical genetics from a Mendelian base has been noted with regret by some geneticists (e.g., Lewontin, 1974; Vogel & Motulsky, 1979, pp. 3, 471). Lewontin charges that it leads psychologists and some geneticists to an unwarranted emphasis on the 'analysis of variance' at the expense of the 'analysis of causes'. The charge is valid, but incomplete: on the reconstruction offered here, psychologists have reasons (although invalid ones) to assume that the analysis of correlation or of variance is the analysis of causes.

Some such reification of statistical models, at least, seems to be implicated in the practice documented in this paper, that of the egregious misinterpretation of purely statistical effects. Such an explanation, however, is conjectural. Whether it is right or wrong, it does not excuse the practice, which must be regarded as indefensible — procedurally, mathematically, and, given the stakes in the I.Q. controversy, socially.

REFERENCES

Galton, F. Regression towards mediocrity in hereditary stature. Journal of the Anthropological Institute, 1886, 15, 246-263.
Scarr-Salapatek, S. Unknowns in the IQ equation. Science, 1971, 174, 1223-1228. (a)
Scarr-Salapatek, S. Race, social class, and I.Q. Science, 1971, 174, 1285-1295. (b)