Psychological Bulletin
1971, Vol. 75, No. 3, 223–224

NOTE ON WHY GENETIC CORRELATIONS ARE NOT SQUARED

ARTHUR R. JENSEN[1]

University of California, Berkeley

Correlations between related persons (twins, siblings, parent-child, etc.) should not be squared in order to determine the proportion of variance they have in common. The correlation coefficient itself is this proportion. We are not using the correlation to predict the variance in a given trait for one set of persons from a knowledge of the trait values of their relatives, but to express the degree of overlap in trait variance, that is, the proportion of variance in common. The rationale of genetic correlations is explained, with examples in terms of a common elements model of correlation.

Psychologists are often puzzled and confused by the fact that geneticists do not square the correlations between twins (or other kinship correlations) in order to obtain the percentage of variance explained by genetic factors. (Or, in the case of correlation between unrelated children reared together, the percentage of variance due to environmental factors.) Recent prominent examples of this confusion are found in Spuhler and Lindzey (1967, p. 403–404) and in Guilford (1967, p. 351–352). These authors incorrectly square kinship correlations and thereby arrive at erroneous conclusions. Most psychologists have learned to treat correlations as the square root of variance explained. But it is incorrect to take the square of twins or other kinship correlations to determine the proportion of variance attributable to genetic or environmental effects. The unsquared correlation itself is correctly interpreted as a proportion. Here is the reason: If the correlation between phenotype (i.e., obtained score) and genotype (i.e., the hypothetical genetic value of the individuals) is r_{pg}, and if the correlation between phenotypes of pairs of individuals with the same genotypes but nothing else in common (e.g., identical twins reared apart in random environments) is $r_{pp'}$, then $r_{pp'} = r_{pg}^2$, or

$$\sqrt{r_{pp'}} = r_{pg}.$$

A good analogy is with test reliability. Two equivalent forms of a test have only their true-score variance in common (analogous to genetic variance) and the error variance (analogous to environmental variance) is not in common, that is, is uncorrelated. The correlation between equivalent forms, r_{tt}, is the reliability, or the percentage of true score variance ("genetic variance") the tests share in common. The $\sqrt{r_{tt}}$ is the correlation of obtained scores with true scores. Thus, the correlation between identical twins reared in uncorrelated environments is directly analogous to the correlation between equivalent forms of a test. The correlation in each case indicates the percentage of variance in common, or the percentage of genetic (or true score) variance.

Another way of regarding the problem is in terms of the "common elements" formula for correlation (given in McNemar, 1949, pp. 117–118). This is

$$r_{xy} = \frac{N_c}{\sqrt{N_x + N_c}\,\sqrt{N_y + N_c}},$$

where

N_c is number of elements common to variables X and Y,
N_x is number of elements unique to X,
N_y is number of elements unique to Y.

A visually simple example is to consider the correlation of half-siblings, who have 25% of their genetic variance in common. The variance can be represented by squares, as in Figure 1. Assume $\sigma_x^2 = \sigma_y^2$, as would be the case for two sets of half-sibs. For simplicity assume σ_x^2, and σ_y^2 each equals 100. (Also, for simplicity assume there is no environmental variance.) Then, applying the common elements formula for correlation, we have

$$r_{xy} = \frac{25}{\sqrt{75 + 25}\,\sqrt{75 + 25}},$$

$$r_{xy} = .25.$$

[1] Requests for reprints should be sent to Arthur R. Jensen, Institute of Human Learning, University of California, Berkeley, California 94720

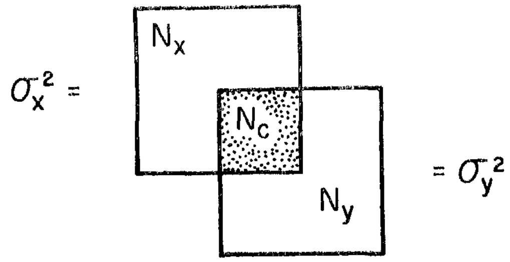

FIG. 1. Correlation of half-siblings who have 25% of their genetic variance in common.

This is the correlation between half-sibs and is also the proportion of the genetic variance they have in common. The correlation between obtained scores and that part of the genetic variance that half-sibs share in common is $\sqrt{.25} = .50$. This can be visualized in Figure 2.

Again, applying the common elements formula:

$$r_{tc} = \frac{25}{\sqrt{75 + 25}\,\sqrt{0 + 25}}.$$

$$r_{tc} = .50.$$

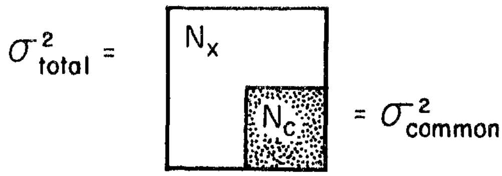

FIG. 2. Correlation between obtained scores and shared genetic variance of half-sibs.

Now, in this case, if we want to know the percentage of total variance that is explained by the common genetic variance, we must square r_{tc}, and this gives .25 or 25%, and, as can be seen in the diagram, this is one-fourth of the total area (variance).

REFERENCES

GUILFORD, J. P. *The nature of human intelligence.* New York: McGraw-Hill, 1967.

MCNEMAR, Q. *Psychological statistics.* New York: Wiley, 1949.

SPUHLER, J. N., & LINDZEY, G. Racial differences in behavior. In J. Hirsch (Ed.), *Behavior-genetic analysis.* New York: McGraw-Hill, 1967.

(Received October 30, 1969)

ERRATA

In the article "Equivalence of Information in Concept Identification" by David Arenberg in the November 1970 issue, the tenth entry in the fifth column of Table 1 on page 358 should read "A_1B_1"; and the eighteenth entry in the fourth column of Table 2 on page 359 should read "A_2D_1."