DOCUMENT RESUME

ED 131 922 PS 008 924

AUTHOR Nichols, Robert C.
TITLE Heredity and Environment: Major Findings from Twin Studies of Ability, Personality and Interests.
PUB DATE 4 Sep 76
NOTE 39p.; Paper presented at the Annual Meeting of the American Psychological Association (84th, Washington, D.C., September 4, 1976)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Academic Aptitude; Aptitude Tests; Birth Rate; *Correlation; *Environmental Influences; Evolution; *Heredity; *Individual Differences; Intelligence; *Nature Nurture Controversy; Personality Development; Personality Tests; Research; Secondary Education; *Twins; Vocational Interests
IDENTIFIERS Eugenics

ABSTRACT

A review of the twin literature and analyses of two large twin samples found identical twin correlations higher than fraternal twin correlations by about .20 for a variety of traits of ability, personality and interests. This was interpreted as indicating that about half of the variation among people in a broad spectrum of psychological traits is due to differences among the people in genetic characteristics. The data also suggest that the environmental influences on ability affect twins reared together in the same way, while the environmental influences on personality and interests affect twins in the same family differently. The implications of these findings for psychological theory, for the nature-nurture controversy, for the explanation of the recent national decline in aptitude test scores, and for future human evolution are discussed. (Author/MS)

Heredity and Environment:
Major Findings from Twin Studies of Ability, Personality and Interests [1/]

Robert C. Nichols

State University of New York at Buffalo

The relative importance of nature and nurture has been one of the oldest and most persistent concerns in the attempt to understand individual differences in human behavior. Experience offers persuasive evidence for both sides of the argument.

On the one hand the observer is struck by the great plasticity and adaptability of human beings to a wide variety of environmental pressures. Consider the variety of primitive societies, each somehow molding the behavior of new generations to perpetuate extremely diverse social patterns. Or within our own culture consider the pervasive effects of sex role, socioeconomic status and racial or ethnic group membership on the behavior of individuals. How can one doubt that the environment molds behavior and that as the twig is bent so is the tree inclined?

On the other hand the observer is struck by the indomitable human spirit rising above environmental adversity to express its inner individuality. The long history of man's inhumanity to man shows that barbarous psychological and physical hardships can be endured and the human personality will again express itself when given half a chance. The human psyche must clearly be made of very tough and durable stuff or else the species could not have survived. Also a long list of anecdotes testifies to the fact that behavioral tendencies run in families much as do skin complexion and body size and shape. How can one doubt that each personality is in many ways a chip off the old block and that blood will tell?

1/ Invited address presented at the American Psychological Association meeting, Washington, D.C., September 4, 1976.

Thus, it seems at the outset that human personality is responsive to the environment, yet has an inherent internal consistency that is resistant to environmental modification. However, as psychologists we would like more detailed and quantitative evidence for this assertion.

There are several lines of evidence that allow the separation of the effects of heredity and environment on behavior. These include studies of the behavioral similarity of relatives, adopted children, separated identical twins, and unrelated persons reared together as well as comparison of the relative similarity of fraternal and identical twins. In studies of animals there are also inbred strain comparisons and selective breeding experiments.

In this paper I will concentrate on the comparison of identical and fraternal twins reared together, however, to my knowledge none of the conclusions to be derived from twin comparisons are contradicted by the other lines of evidence.

The Twin Method

Twins form a natural experiment that among Caucasian Americans occurs at a rate of about one set of twins in every 100 births. About a third of these are identical twins with identical genetic material, and two-thirds are fraternal twins with the same genetic relationship as ordinary siblings, (i.e., about half their genes in common). Since half the fraternal twins are of opposite sex, there are about equal numbers of identical and same-sex fraternal twins, which are the groups usually studied. The zygosity of twins, whether they are identical or fraternal, can be diagnosed with good accuracy from readily observable physical characteristics such as height, the color, texture and pattern of growth of hair, eye color, ear lobe attachment, facial physiognomy, etc. If the twins of a set are alike on all of these genetically determined features the probability is high that they are identical twins. If they are definitely different on one or more, they are fraternal. An added degree of precision can be achieved by also considering a number of independently inherited blood types.

Twins raised together in the same family are exposed to pretty much the same environment whether they are identical or fraternal, thus, if identical twins are more alike on some psychological trait than are fraternal twins, this may, with some assumptions, be attributed to their greater genetic similarity.

Most psychological traits are continuously variable, rather than categorical, and the index of twin similarity that is used with such variables is the intraclass correlation, which in a large sample is usually numerically identical to the product moment correlation between twins. The intraclass correlation, like a reliability coefficient, is the proportion of the variance of the trait in the sample that is common to the twins of a set. The major results of a twin study, then, are the intraclass correlations for identical and fraternal twins for each trait that is studied.

Since identical twins have all their genes in common and fraternal twins have half their genes in common, the difference between the intraclass correlations for identical and fraternal twins is the proportion of variance attributable to half the genetic influence on the trait. Thus, if we double the difference between the intraclass correlations, it would seem that we have the heritability, the proportion of variance in the trait attributable to heredity.

There are, of course, a number of complicating factors that can distort this seductive simplicity, and these will be discussed later in the paper.

In addition to the degree of hereditary influence on a trait that is indicated by the difference between the correlations, the size of the correlations can also tell us something about the nature of the environmental influences on the trait. If the salient environmental influences vary from family to family but tend to be the same for the children within a family--such as characteristics of the parents, the home, and the community--the correlations will tend to be relatively higher. This is because the common environmental influence on both twins of a set will tend to make them similar and thus increase the correlations for both identical and fraternal twins. On the other hand, if the salient environmental influences affect the twins within the same family differently the correlations for both identical and fraternal twins will tend to be relatively lower.

The Twin Literature

Keeping these thoughts in mind let us now look at the sort of twin correlations that are typically found. Since Galton's pioneering twin study in 1875 there have been over 100 published studies comparing the relative similarity of identical and fraternal twins on a great variety of psychological traits. A former student of mine, Dr. Nancy Breland, reviewed this twin literature up to 1971 and extracted 756 pairs of intraclass correlations, 2/ (Breland, 1972). With some advice from colleagues, Nancy and I sorted the correlations according to the trait measured into the broad domains of ability, personality and interest. Within each domain we further classified them according to the specific trait. In those instances in which a test was used that was not known to us or in which the trait could not be unambiguously grouped with other studies of the same trait, the correlations were temporarily discarded. This provided a way of organizing a large body of data, although tests with

---

2/ A list of these correlations with references to the articles from which they were obtained is available from the author on request. The correlations were recorded as presented in the article. About 20 percent had been corrected for attenuation; the rest were observed correlations.

similar names, which we grouped together, were no doubt in many instances measuring quite different traits.

The results for the ability domain are shown in Figure 1. This figure shows considerable variation among studies, which is a fact of life that psychologists have had to learn to live with, but there is also a striking overall consistency. The correlations were predominantly high and positive, showing that twins tend to be quite alike in a variety of abilities. Moreover, identical twins tended to be more similar than were fraternal twins on all 12 traits of ability. If we consider the weighted averages represented by arrows to be one large composite twin study based on several thousand sets of twins of each kind, the difference between identical and fraternal intraclass correlations ranged from .25 for general intelligence to .11 for divergent thinking.

Figure 2 shows the analogous results for the interest domain, and the picture is quite similar to that for abilities, except that the correlations were somewhat lower. The difference between weighted averages for identical and fraternal twins ranged from .22 for artistic interests to .11 for business or enterprising interests.

Figure 3 shows analogous results for the personality domain, which are similar to those for interests except that the horizontal lines tend to be somewhat longer, indicating greater variation among studies. The difference between weighted averages for identical and fraternal twins ranged from .27 for extroversion to .19 for masculinity-femininity.

Confronted with the remarkable similarity of results for ability, interests and personality as well as for the more specific traits within the three domains, one might next ask whether this survey of the twin literature reveals any significant difference between traits or between domains.

The weighted composite results of all twin studies are not appropriate for answering this question because the different traits were investigated by different studies using different samples of twins. Thus, the large differences we have noted between different studies of the same trait could produce spurious differences between composite results for different traits. The best evidence we have on this issue then, comes from considering each study as an independent attempt to estimate population values for a given trait. We can then ask whether studies of different traits produce results that cannot reasonably be attributed to the variation among studies that is observed when they investigate the same trait. For this analysis each study was an equal unit and was not weighted by its sample size.

Table 1 shows the unweighted mean correlations for the various traits that were shown in the previous three figures. The basic data are the same as in the figures, but the results are slightly different because the studies were given equal weight regardless of sample size. We can now ask whether the observed variation among traits in the average difference between identical and fraternal correlations can be attributed to chance. A one-way analysis of variance showed the differences among the traits in the ability domain to be not significant at the five percent level. Similar analyses in the interest and personality domains showed even greater likelihood that the observed differences among traits are due to chance. Thus, the difference in similarity between the two kinds of twins does not differ significantly among traits within the three domains. We are now back to the mean difference between identical and fraternal correlations for the three domains, which for ability, personality and interests respectively were .21, .19 and .18. These three numbers are clearly not significantly different from one another. Thus, we must conclude that a twin study is likely to find a difference between identical and fraternal correlations of about .20 regardless of the domain or the trait that is being investigated!

The tendency for both the identical and fraternal correlations to be higher in the ability domain than in the personality and interest domains was highly significant statistically. The average correlation involving an ability measure was higher by about .25 than that involving an interest or personality measure.

Without attempting to interpret these correlations precisely at this point, I will merely state what seems their most obvious implication: individual differences in all traits of behavior, from general intelligence to fingernail biting, are due in roughly equal parts to genetic differences and to environmental differences. The environmental factors that influence abilities tend to affect members of the same family in the same way, while the environmental factors that influence personality and interests tend to affect members of the same family differently.

The National Merit Twin Study

I would now like to discuss some new results from twin data that I was mainly responsible for collecting. 3/ Although the data were collected over ten years ago, some of the results are, in fact, new. Personal and organizational factors delayed the analysis, and the book containing a fairly complete report of the entire study was published only a few months ago (Loehlin and Nichols, 1976).

During the 1960's the National Merit Scholarship Corporation conducted an annual, nationwide testing program in which each spring a three hour test of educational development was given to selected eleventh grade students in most high schools in the United States. Thanks to the support for research provided by John M. Stalnaker, then President of National Merit, we were able to ask on

---

3/ These data were collected as part of the research program of the National Merit Scholarship Corporation, which was supported by grants from the Ford Foundation, the Carnegie Corporation of New York, and the National Science Foundation.

the test form whether or not each of the roughly 600,000 students who took the test in 1962 was a twin.

By pairing twins who attended the same school and who had the same last name and home address we identified about 1500 sets of same-sex twins. These twins were each sent a questionnaire asking for detailed reports on a number of hereditary physical characteristics, which were used for diagnosis of zygosity. Subsequently, blood samples were obtained from 124 sets of these twins, and the questionnaire diagnosis agreed with the diagnosis based on extensive blood typing in 93 percent of the cases (Nichols and Bilbro, 1966). Usable questionnaires were received from about 1200 sets.

These twins were then each sent an additional packet of questionnaires concerning their behavior, attitudes, goals, interests and personality and a separate questionnaire to be completed by a parent. The packet contained the California Psychological Inventory and a long questionnaire developed specifically for the study that together required about three hours for each twin to answer. Complete data were obtained from 850 sets of twins, of which 60 percent were diagnosed identical and 42 percent were male.

This procedure yielded data on a large number of sets of same-sex twins, all about the same age with each set raised together in the same family. These twins are, of course, not representative of any specific group to which statistical generalizations might be made, however, comparison with available norms suggests that they are not particularly unusual in any respect except for being twins and for being National Merit test takers. Like other National Merit test participants, they averaged about one standard deviation above high school students in general on measures of scholastic aptitude. They showed about the same variability as students in general on tests of ability, personality and interests. Except for the ability tests all data were obtained via mailed questionnaires. Checks of internal consistency of the questionnaire responses and comparison of reports of the same information by the two twins and their parent indicate that the questionnaires were carefully completed. The data appear to be about the same quality as is usually obtained in group testing of college students.

Three years later, in 1965, a second twin sample was obtained. The twins among the almost 800,000 participants in the 1965 National Merit testing program were identified as part of a broader question about birth order. The same-sex twins from this testing were sent a revised form of the physical similarity questionnaire from which 1300 identical sets and 864 fraternal sets of twins were diagnosed. About two years later David P. Campbell, then at the University of Minnesota, mailed these twins the form of the Strong Vocational Interest Blank appropriate for their sex. Usable Strong tests were obtained from 669 male and 949 female sets of twins of which 61 percent were diagnosed identical. As in the 1962 sample, females and identical twins were somewhat more cooperative in providing data than were males and fraternal twins.

The diverse data on these two large twin samples lend themsleves to a number of different analyses, but here we will limit our attention to the relative similarity of identical and fraternal twins on the various measures. All told, data were available on the 1962 sample for over 1600 test scores and questionnaire items, and the computer obediently spewed out intraclass correlations for all of them.

Differences Among Traits

When the correlations for the major scores representing ability, personality and interest are plotted on the figures showing the results of past studies they blend right in, and the present investigation could well serve as the modal twin study. Our attention was then directed to a more detailed investigation of the striking implication of the literature that the difference between identical and fraternal correlations, and thus the heritability, is about the same for all psychological traits.

To study this question, John Loehlin performed a series of analyses that took advantage of the relatively large sample of twins and the diversity of variables in this study. John's innovative method was to compute correlations for a variety of variables separately in random halves of the twins of each sex and to look for agreement between random halves and between sexes in the rank order of differences between the correlations for the two kinds of twins. In this way, any tendency for some traits to show consistently larger differences between identical and fraternal correlations than do other traits could be detected. Traits with high heritabilities should show large differences and those with low heritabilities should show small differences in both random half-samples.

This method cannot simply be applied to the entire list of 1600 variables, because of the large differences among them in reliability. Unreliable variables would tend to have consistently low correlations with correspondingly small differences between them simply because of their low reliability.

The first application of the random-half method, then, was to the 18 standard scales of the California Psychological Inventory, which do not differ greatly in reliability. The scales were ranked in terms of the size of the difference between identical and fraternal correlations in random half-samples of each sex. There was no agreement in these ranks between sexes or between random half-samples. The average Spearman rank correlation was -.05. There was also no agreement between the rank order of the CPI scales in this study with that reported in a previous twin study of the CPI conducted by Irving Gottesman (Gottesman, 1966; Nichols, 1966).

Using the 1965 sample, a similar random-half analysis was done using 88 Strong Vocational Interest Blank scales for males and 69 scales for females. Again there was no agreement between random halves in the rank order of the scales in terms of the difference between identical and fraternal correlations.

There was also no agreement between the rank order of identical-fraternal differences for the five subtests of the National Merit Scholarship Qualifying Test either for the two sexes or for the 1962 and 1965 samples. However, this may not be especially surprising, since these subtests are highly intercorrelated.

To give any differences among traits the maximum chance to show themselves, one should use as diverse a set of variables as possible. For this purpose John Loehlin performed a series of cluster analyses on all of the 1500 or so questionnaire items available on the 1962 sample to develop a set of diverse clusters, each with reasonable internal consistency. This process yielded 70 clusters in which no variable was in more than one cluster and every variable in a cluster correlated at least .30 with every other. The number of variables in the clusters ranged from 3 to 13 with a median of 4. Although the clusters were formed on an entirely statistical basis, in almost all cases the content was fairly homogeneous and readily interpretable. There was great diversity of content among the 70 clusters, which included abilities, interests, life goals, self-concept ratings, ideal-self ratings, activities, descriptive adjectives, physical complaints, attitudes and CPI items.

Differences between identical and fraternal intraclass correlations were computed for the 70 clusters for random half-samples of males and females. There was no agreement between sexes or random half-samples in the rank order of these differences. The average Spearman rank order correlation was .02.

Thus, it seems that it is quite difficult to find evidence of greater genetic involvement with some psychological traits than with others, even with the relatively large sample of twins available for this study.

In a final attempt to find such evidence we looked at differences between identical and fraternal twin correlations of individual CPI items in random half-samples. To avoid the difficulties presented by low item reliability, we selected only those CPI items that could stand alone psychometrically. Goldberg and Rorer (1964) obtained 3- to 4-week test-retest data for the CPI item pool for three samples of college students ranging in size from 95 to 179. We retained only those items that had test-retest reliability coefficients of at least .50 in each of the three samples. There were 179 such highly dependable items.

Next, we sorted out from among these reliable items those in which the intraclass correlation between identical twins was at least .20 higher than that between fraternal twins ("high-difference" items) and those in which either the correlation between identicals was no more than .02 above that between fraternals or the fraternal correlation was higher ("low-difference" items). This procedure was carried out separately in the two random halves of the total sample. The question was simply: "Are high and low identical-fraternal differences consistent properties of particular items, or are we screening chance sampling fluctuation?"

There were 55 and 54 items meeting the criterion of high difference in the two half-samples and 38 and 31 items meeting the criterion of low difference. There was a significant tendency for the low-difference items to be the same in the two half-samples, however, there was no such tendency for the high-difference items. Eleven items had low differences in both half samples, and only 6.6 would be expected by chance.

Among the eleven items showing a low difference in both half samples were several expressing social attitudes, and this content did not occur among the high-difference items. These items were: "A person who doesn't vote is not a good citizen," "I do not like to see people carelessly dressed," "I believe women should have as much sexual freedom as men," and "People have a real duty to take care of their aged parents, even if it means making some pretty big sacrifices."

With this hint we noticed that elsewhere in the questionnaire individual items concerned with attitudes toward racial integration and federal welfare programs and with belief in God also showed low differences.

Thus, there is some evidence that specific social attitudes are less dependent on the genes than are most other psychological traits. It is somewhat reassuring to find that identical twins were not consistently more similar than were fraternal twins on everything. Otherwise we might have to entertain a hypothesis about some special ESP or perhaps colusion on questionnaires among identical twins. In this vein we might note with some feeling of relief that the identical twins were not noticeably more alike than the fraternals on such items as reports of the size of their high school class, the size and urbanization of their home towns or the presence of various items in their home.

Although there were practically no dependable differences among psychological traits in the difference between identical and fraternal correlations, the size of the correlations did differ reliably among trait domains. As in previous studies, correlations tended to be higher for abilities than for personality and interests.

<u>Interpretation of Twin Correlations</u>

Table 2 shows the median correlations obtained for several major groups of variables. A random half-sample analysis showed that the difference between identical and fraternal correlations for the various classes of variables was not dependably different, although the varying size of the correlations, (e.g., the mean correlation for the two kinds of twins) was dependable. We may now attempt to interpret these correlations in more detail.

First we should adjust the correlations for two known sources of error in the study--the variables were not measured with complete reliability and the zygosity of the twins was not diagnosed with complete accuracy. Table 3 shows the correlations corrected for these attenuating influences. The reliability estimates used for these corrections are shown in the first column of the table. The best estimate from the blood studies mentioned above is that about seven percent of the twins of each kind were misdiagnosed, and the correlations were adjusted for the effect of these errors in diagnosis of zygosity. The effect of both of these corrections was to increase the observed correlations. The difference between the corrected identical and fraternal correlations is now about .30, which implies a heritability of about .60. Because the heritability estimate is subject to sampling fluctuation and is fairly sensitive to the estimate of the reliability of the test, we should probably not state this more precisely than to say that about half the variance in these traits seems to be attributable to genetic differences.

Additional correction for assortative mating would not change very much the heritability estimate for personality and interest measures, where quite low positive correlations between spouses are typically found. However, husband-wife correlations on the order of .40 to .50 are generally reported for general intelligence (Vandenberg, 1972), and allowance for this would increase the heritability estimate for abilities to about .70. [4]

4/ The correction for unreliability of measurement is

$$r_c = \frac{r_o}{r_{tt}} ,$$

where r_c is the corrected correlation, r_o is the observed correlation and r_{tt} is the reliability coefficient. (Footnote continues on page 15.)

(continuation of footnote from previous page)

The correction for misdiagnosis is

$$r_{cMZ} = \frac{r_{MZ} - (r_{DZ}\, e_{MZ})}{1 - e_{MZ}},$$

where r_{cMZ} is the corrected correlation for identical twins, r_{MZ} is the correlation for identical twins (corrected for attenuation), r_{DZ} is the correlation for fraternal twins (corrected for attenuation), and e_{MZ} is the proportion of fraternal twins that are erronenously diagnosed as identical. The comparable formula for the fraternal correlation may be obtained by reversing the MZ and DZ subscripts.

The heritability (h^2) is

$$h^2 = \frac{r_{MZ} - r_{DZ}}{1 - r_g},$$

where r_g is the genetic correlation between fraternal twins, which is assumed to be .50 if there is no assortative mating. With a correlation between spouses of .50 and a heritability of .70, r_g becomes .57 (Jensen, 1967).

There are additional qualifications that should be placed on heritability calculated from twin correlations. Non-additive genetic effects (dominance and epistasis) are included in the heritability figure. Thus, it is often described as "heritability in the broad sense," the total genetic effect, in contrast to "heritability in the narrow sense," which is the heritability that would be realized in selective breeding. Variance attributable to the correlation of genetic and environmental influences is also included in the heritability figure. This correlation might be either positive (those with the more favorable genes are exposed to the more favorable environment), negative or curvilinear (those genetically extreme on a trait are influenced by the environment to be less extreme). Other complications, such as differences in the intrauterine environment for the two kinds of twins, have also been suggested.

Some observers have cautioned that the greater behavioral similarity of identical twins may be due in part to a greater similarity of their environment rather than of their genes. Reports by the twins and their parents indicate that identical twins do in fact spend more time together, have more similar early experiences, and are treated somewhat more alike by parents than are fraternal twins. However, this does not seem to be a reasonable explanation for their greater psychological similarity. Within twins of the same type, greater similarity of experience was not associated with greater similarity on the psychological traits with which we are concerned. In other words, the difference in similarity of environment that has been noted for the two kinds of twins does not appear to result in corresponding differences in psychological similarity. Thus, the best explanation for the twin data in our study and in the literature is that about half of the variation among people in a broad spectrum of psychological traits is due to differences among the people in genetic characteristics.

There is at least one theory which suggests that under long-term evolutionary conditions one might expect traits to tend toward roughly equal (and moderate) heritabilities. ~~The theory derives from arguments~~ outlined by Allen (1970). It holds that, if the heritability of a trait is low, gene mutations affecting the trait will tend to accumulate, increasing its genetic variance. Once the genetic variance becomes large enough relative to environmental variation so that the heritability of the trait is appreciable, stabilizing natural selection will begin to operate on the trait to slow and eventually to stop further increase in its genetic variability and hence to hold heritability at a stable level. If relevant environmental variation were to decrease, the trait heritability would temporarily rise, permitting selection to act more effectively on the genetic variation of the trait, bringing the genetic variation (and thus the heritability) back down again. Generally speaking, then, under this hypothesis all traits tend toward moderate levels of heritability because the genetic component of variation of any trait tends to increase until the process of natural selection can "see it" against the background of environmental variation present and hold it stable. This suggests that differences in the importance of different traits for reproductive fitness will principally be reflected in the total amount of variation present, rather than in the relative proportions of this variation that are genetic and environmental. A trait that is critically important for reproduction will show little variation among individuals and a trivial trait will show a lot, but their heritabilities will be about the same.5/

For this mechanism to work, the general level of environmental influence on any given trait must remain fairly constant on the scale of tens of thousands or hundreds of thousands of years on which human biological evolution takes place. The specific environmental influences need not always be the same, but their general level of impact must remain fairly constant. On the face of it, this does not seem very reasonable. The tremendous changes in the human environment that industrialization has produced over the past several hundred years must certainly have changed the environmental impact on human behavior, if only by reducing the privations and noxious circumstances that would seem to characterize life in the wild. But we must remember that we do not know what the major environmental events that produce differences in human personality are. If the critical events are prenatal biological factors, or basic parent-child emotional relationships, or the like, or if they occur in some relatively short critical period it may be reasonable to assume that their

5/ Perhaps this is why sexual libido and love of children are so widespread while love of statistics is only as strong in a few of us.

impact has remained fairly constant over the millenia.[6/]

It has been customary for those discussing the heritability of human behavior to point out that the heritability coefficient is a population statistic that is specific to a given group at a given time, and it has been suggested that heritability may vary widely even among sub-cultures in the United States. The current line of argument, on the other hand, implies that the genetic and environmental factors responsible for individual differences are rather basic properties of the human condition, and that one would expect to find roughly similar heritabilities over a fairly wide range of circumstances.

Environmental Correlation

Another perspective on the character of relevant environmental factors may be obtained by considering the different levels of twin correlation prevailing in different trait domains. Assuming that the degree of genetic influence causing twins to be alike is roughly the same for all trait domains, the differences in the level of correlation (the average of the correlations for both kinds of twins) in the various domains can be attributed to differences in the similarity of environmental influences on the twins of a set. Thus, although we do not know specifically what the environmental factors are, we can say something about the degree to which they affect twins raised together in the same family in the same way. More precisely, if we know the heritability, we can calculate the correlation between twins of the salient environmental influences

---

6/ If any trait has a higher heritability than most others it would seem from the accumulated evidence, to be intelligence. Since selection for a trait reduces heritability, it has been argued that its high heritability indicates that intelligence, as currently measured, has not been subject to strong natural selection in man's evolutionary past (McClearn and DeFries, 1973, p. 254). A more likely explanation is that the variation in environmental influences on intelligence has been greatly reduced in recent times.

that are implied by the intraclass correlations.[7/] These environmental correlations for the various domains are shown in Table 4. Separate estimates of the same environmental correlation were obtained from the intraclass correlations for identical and fraternal twins. The third column in the Table shows the average of these two estimates. These environmental correlations indicate the degree to which the environmental influences that produce individual differences in the trait, whatever they may be, have the same effect on both twins of a set.

These environmental correlations are subject to sampling fluctuation, as is shown by the different estimates obtained from the two kinds of twins, and they are also somewhat sensitive to the estimate of reliability that was used in correcting the correlations for attenuation. Thus, small differences between traits should not be taken too seriously. There was a very clear tendency, however, for abilities and activities to have high environmental correlations and for personality and interests to have low environmental correlations, a finding that is consistent with previous twin studies.

It is not difficult to accept the high environmental correlation for abilities and activities, since one might reasonably expect that the relevant environmental inputs would be associated with the characteristics of the parents,

---

7/ Under certain simple assumptions, the pheonotypic (observed) correlation (r_p) is the sum of the genetic (r_g) and environmental (r_e) correlations, weighted by the heritability (h^2):

$$r_p = h^2 r_g + (1 - h^2) r_e,$$

and therefore,

$$r_e = \frac{r_p - h^2 r_g}{1 - h^2}.$$

the home, the school and the community, all of which are the same for both twins of a set.

But what about the very low or even negative correlation between twins in the environmental influences on personality and interests? Can this possibly be true? Surely such factors as the parents' child-rearing philosophies, and the psychological ambience of the community and the home have some influence on the development of personality; and these things are the same for twins reared together. In fact, almost all of the environmental antecedents of individual differences in personality that have been suggested by psychologists and others are the same for twins reared together.

One possible explanation of this paradox lies in the special environmental situation of twins and in our reliance on self-report measures of personality. Each twin has the other twin as a major part of his environment, and this may lead to competition or to contrast effects between them. If a twin's reference point for self-definition is the other twin, and if others around him are continually contrasting the pair members, it seems plausible that they might end up seeing themselves as much less similar in personality and interests than they actually are. Since our personality and interest measures were all based on some form of self report, such a contrast effect could mask the similarity produced by the common environment. This hypothesis obtains some support from the fact that the somewhat indirect self-report measures, such as the CPI and Strong scales, show low positive environmental correlations, while the more direct ratings of self concept show negative environmental correlations.

There are, however, at least three arguments against this explanation. 1) Some twelve twin studies in the literature have used non-self-report-measures of personality such as hypnotic susceptibility, musical preferences, flicker fusion, auto-kinetic movement, speed of decision, free association, social intelligence, and color-form movement. The unweighted average intraclass correlations for the more objective personality measures from these studies were .47 and .30. These correlations show the usual difference between correlations for identical and fraternal twins near .20, ~~but~~ yield near zero estimates of environmental correlation. 2) One would expect that the contrast effect would vary in some systematic way across personality traits and for the two sexes, but no such pattern was observed. 3) The degree to which twins of the same kind tended to associate with each other was unrelated to personality differences between the twins, although one might expect a strong contrast effect to be sensitive to the amount of contact between the twins.

Another possible explanation for the low environmental correlation is that environmental influences on personality are actually highly variable situational factors. If the ways in which environment affects personality are sufficiently complex, contingent and subtle, they could appear random in their effects on twins within a family, which is how they, in fact, do seem to appear.

The present state of psychological research is such that it is difficult to find uncontested evidence showing substantial, enduring effects of major environmental variables on personality and interests that can counter the impression from twin studies that such environmental inputs are highly specific, if not actually trivial, events. I have also not been able to find good studies of the similarity in personality of unrelated children raised together and of twins or ordinary siblings raised apart, which might help resolve the issue. Studies of family resemblance have in the past been mainly the tool of those interested in genetic influences. However, resolution of the current paradox might throw new light on the nature of environmental influences as well.

Implications

Having assembled and reported all of these data, tradition now entitles me to make some considerably less rigorous comments on their broader implications. From the many implications that deserve comment I will select two.

First, if we are to understand human individual differences and perhaps aspire to predict and control them, we must not ignore the genetic mechanisms. Psychologists have tended in recent years to prefer environmental explanations for individual differences, and most of our research and applied attempts to change behavior have been directed toward the environment. However, the data seem to indicate that this excessive environmentalism is ignoring at least half of the problem. Thus, I suggest that increases in understanding of individual differences will be more rapid if genetic hypotheses receive roughly equal time.

To take a concrete example, let us look at attempts to explain the recent decline in average student aptitude test scores. The average score on the Scholastic Aptitude Test and the American College Testing Program test has been declining steadily for over ten years, and the accumulated decline is substantial--a quarter of a standard deviation or more (Munday, 1976). Public attention was first attracted to the decline in scores on the SAT, but subsequent reports have shown that the lower average scores are very widespread. They are found in most state testing programs in elementary and secondary schools, in both suburban and inner city areas (Armbruster, 1975). A decline has also been observed in some tests of the National Assessment of Educational Progress and in Canada (Munday, 1976).

The decline, which apparently cannot be accounted for by artifacts of the tests or by changes in the pool of students tested, has already had important practical consequences: Colleges have lowered their admissions standards and remedial programs at the college level are on the increase.[8/] But there is a more basic psychological question. Does the score decline represent a-

---

8/ New York Times, March 7, 1976.

national decline in intelligence? If so, what is causing it, how long is it likely to continue, and what, if anything, can be done to change it?

A great variety of explanatory factors have been proposed. These include characteristics of the schools, such as changed curricular emphasis and increased permissiveness; characteristics of the society, such as television and the increasing use of drugs; and characteristics of the home such as changing sibling configurations and the increase in working mothers (Harnischfeger and Wiley, 1976; Zajonc, 1976). These explanations have two things in common: 1) they are all environmental, and 2) they are, with the exception of sibling configuration, unaccompanied by evidence that differences in the proposed explanatory factor are related to individual differences in intelligence. This is a good example of the current environmentalist bias--a tendency to prefer environmental explanations and to argue persuasively for them in the absence of strong evidence. I am sure that some researchers have considered genetic explanations, but I have not encountered them in the sample of the psychological and educational literature or in the popular articles that I have read.

There is, however, a long history of research on intelligence and family size that has been largely ignored in discussions of the current decline in test scores. One reason for this oversight may be that the decline has been observed and discussed primarily in an educational context, and, as a result, it has often not been seen as a decline in intelligence. Most discussions have been in terms of operational euphemisims such as academic performance, reading skills, mathematical ability, vocabulary level, and the like, but the large common element in all of these measures is what psychologists call intelligence.

Intelligence likely has the highest heritability of any psychological trait, and it should, therefore, be sensitive to selective mating. A correlation between intelligence and number of siblings on the order of -.20 has been reported by a number of investigators in the United States and Europe over the last 50 years or so, and a similar result was found in the National Merit testing from which the 1965 sample of twins was selected. Thus, those families that tend to produce higher scoring children tend to produce fewer of them and those that tend to produce lower scoring children tend to produce more of them. From this, one would expect average test scores to decline, yet up until about 1960 average scores were increasing. It was argued that urbanization, the lengthening of formal education and increased penetration of the communication media were masking a genetic decline. However, a more widely accepted explanation of the paradoxically rising scores came at about the time the scores were reaching their peak. Articles published in 1962 and 1963 reported that when a total population was studied, including those with childless marriages and individuals never marrying, there was actually a slight positive correlation between intelligence and number of children produced (Bajema, 1963; Higgins, et al, 1962; Waller, 1971).

This positive correlation between intelligence and fertility [9/] was observed in samples of several hundred people in Minnesota and in Michigan whose families were complete before the onset of the current decline. Since then we have seen the post-war baby boom, the advent of oral contraceptives and legal abortion, and the current declining birth rates. With these upheavals in birth patterns it is possible that the correlation of intelligence with fertility has changed from positive to negative, and is the major factor responsible for the current decline of average test scores. [10/]

9/ A biologist reviewer advised that "fecundity" is the preferred term for number of children produced, since "fertility" implies the capacity to have children. "Fertility" is used here for consistency with usage in the literature and by the U.S. Census.

10/ Another genetic mechanism that probably affects intelligence is heretosis, the increase observed in traits that have been subject to evolutionary selection when different strains are mated. The panmixia resulting from
(footnote 10 continued on page 25)

To test this hypothesis we would like to have recent data on intelligence and number of children produced for a representative sample of the population. These data are not currently available, however, the U.S. census publishes data showing education and number of children. For group comparisons years-of-education-completed can serve as an indicator of intelligence, and, in its own right, education-of-mother is substantially correlated with measured-intelligence-of-child whether or not the mother actually raised the child (Honzik, 1957). Of the age groups used in the census reports, the range 35 to 44 is most appropriate for our purposes, since the families of women this age are practically complete and typically contain children in their late teens where the score decline has been most clearly observed. Table 5 shows birth rates by education for women of this age for 1960, a year near the test score peak, and for 1974, the latest year available. In 1960 there was a pronounced tendency for women with more education to have fewer children than did those with less education. Birth rates increased for all educational levels from 1960 to 1974, reflecting the higher post-war birth rates, but the largest increases were among the lower educational levels. With the exception of the relatively small groups with eight or less years of education, the change in birth rate decreased monotonically with increasing education. Thus, the negative correlation of education with fertility observed in 1960 was more pronounced in 1974. This finding tends to support the proposed genetic explanation for the recent test-score decline.

10/ the increased immigration of diverse groups into the U.S. around the turn of the century may be largely responsible for the large increase in intelligence that occured here between the two world wars (Tuddenham, 1948). Since heterosis is the cause of the increase in average height observed in recent times, it may be possible to use changes in height as an indicator of the heterosis effects to be expected in intelligence.

The second implication that I will mention concerns the possibility of changing human nature by manipulating the human gene pool. The substantial heritabilities of almost all psychological traits suggest that such a program could be successful. When biologists talk about such possibilities they usually blithely assume that almost any psychological trait is almost completely dependent on the genes, but psychologists have usually been skeptical of this assumption. If a new individual could somehow be produced that was an exact genetic copy of, say, Albert Einstein, would he grow up to be like the Einstein we admire or would he be someone else? The data suggest that he would, indeed, behave much like the former Einstein, but that he would also have a number of idiosyncrasies of his own.

Man's character has, up until now, been shaped by the capricious forces of nature, but he is about to break out into a new era in which he will intelligently control his own future evolution. Or perhaps we should just say that he will control it without characterizing the control as necessarily intelligent. Hopefully, the current rapid growth of the human population will be limited by rational human decisions rather than by natural catastrophe, and these decisions will determine the future evolution of man. There will be evolutionary changes and the most far-reaching of these will be changes in human psychology.

While we are still on the threshold, psychologists should be considering what might be the most desirable mix of psychological traits and what might be the most humane methods of moving mankind in the desired direction. This is not an easy task. The problem of values has been the major stumbling block of engenics programs in the past, and it will become much more acute as population pressures force intrusions into individual reproductive decisions and as biological technology develops frighteningly powerful techniques of genetic manipulation.

Perhaps the most reasonable proposal for a beginning in the control of human evolution is the method of germinal choice that was proposed by the late Nobel-Prize-winning biologist, Hermann M. Muller (Muller, 1965). In this method sperm is collected from selected outstanding men and is kept safely stored at very low temperatures. After time permits an objective evaluation of the life of the donor, his sperm is made available to women who desire to conceive a child with it. Muller reasoned that this opportunity for genetic choice would be immediately adopted by those women now relying on more haphazard methods of artificial insemination, and, as its success was demonstrated, many couples would opt for the chance to raise an exceptional child. Muller suggested that, in addition to the absence of genetic disease, the traits to be used in selecting donors be intelligence and cooperativeness on the ground that these traits have been most responsible for successful human evolution in the past.

Rapid progress in biological technology has already extended the potentialities of the method of germinal choice beyond those discussed by Muller. It is now technically feasible to collect ova as well as sperm from selected donors and to implant the fertilized ova in the uterus of the recipient woman (Glass, 1972). With this form of "prenatal adoption" 11/ genetic selection would be much more effective.

Even more powerful biological techniques are on the horizon. Cloning has been accomplished in certain amphibians, (Handler, 1970). The haploid nucleus can be removed from an unfertilized frog egg and replaced by the dip-

11/ By use of the term "prenatal adoption" Glass (1972) intended to convey the idea that traditional family values would be preserved. He based this assertion on the testable psychological hypothesis that parental emotions and family cohesiveness result from the experience of the growth of the fetus within the mother and its subsequent birth rather than from knowledge that the child carries the parental genes.

loid nucleus of a body cell from a frog embryo. Such an egg can develop into a normal frog with the same genetic constitution as the frog embryo whose body cell provided the transplanted nucleus. Thus, a clone, a number of genetically identical individuals, can be produced. Although the technical problems are formidable, there seems to be no reason in principle why the process cannot be extended to man.

The means for genetic change are at hand and much more effective means will likely become available in the not-too-distant future. Thus, the changes in human intelligence and personality that psychologists, educators and others have been ineffectually trying to bring about by environmental manipulation can probably be much more effectively accomplished by genetic manipulation. As Bentley Glass (1972) has suggested, perhaps it would be wise to pause on the threshold long enough to consider which way we should go. However, the world moves on, and changes will occur, willy-nilly. To decide to do nothing is just as fateful a choice as to decide to do something, and it should be as carefully considered.

References

Allen, G. Within and between group variation expected in human behavioral characters. Behavior Genetics, 1970, 1, 175-194.

Armbruster, F. The U.S. primary and secondary education process. Croton-on-Hudson, N.Y.: Hudson Institute, 1975.

Bajema, C. Estimation of the direction and intensity of natural selection in relation to human intelligence by means of the intrinsic rate of natural increase. Eugenics Quarterly, 1963, 10, 175-187.

Breland, N. S. A new approach to estimates of heritability from twin data. Unpublished Research Qualifying Paper, Department of Educational Psychology, State University of New York, Buffalo, N.Y., 1972.

Glass, B. Eugenic implications of the new reproductive technologies. Social Biology, 1972, 19, 326-336.

Goldberg, L. R., and Rorer, L. G. Test-retest item statistics for the California Psychological Inventory. Oregon Research Institute Research Monographs, Vol. 4, No. 1. Eugene, Ore.: Oregon Research Institute, 1964.

Gottesman, I. I. Genetic variance in adaptive personality traits. Journal of Child Psychology and Psychiatry, 1966, 7, 199-208.

Handler, P. (Ed.) Biology and the future of man. New York: Oxford University Press, 1970.

Harnischfeger, A., and Wiley, D. E. Achievement test scores drop: So what? Educational Researcher, 1976, 5, 5-12.

Higgins, J. V., Reed, E. W., and Reed, S. C. Intelligence and family size: A paradox resolved. Eugenics Quarterly, 1962, 9, 84-90.

Honzik, M. P. Developmental studies of parent-child resemblance in intelligence. Child Development, 1957, 28, 215-228.

Jensen, A. R. Estimation of the limits of heritability of traits by comparison of monozygotic and dizygotic twins. Proceedings of the National Academy of Sciences, 1967, 58, 149-156.

Loehlin, J. C., and Nichols, R. C. Heredity, environment and personality: A study of 850 sets of twins. Austin, Texas: University of Texas Press, 1965.

McClearn, G. E., and DeFries, J. C. Introduction to Behavioral Genetics. San Francisco: W. H. Freeman, 1973.

Muller, H. J. Means and aims in human genetic betterment. In Sonneborn T. M. (Ed.) The control of human heredity and evolution. New York: Macmillan, 1965. Pp. 100-122.

References (cont.)

Munday, L. A. Declining admissions test scores. ACT Research Report, No. 71. Iowa City, Iowa: American College Testing Program, 1976.

Nichols, R. C. The resemblance of twins in personality and interests. National Merit Scholarship Corporation Research Reports, Vol. 2, No. 8. (Reprinted in Manosevitz, M., Lindzey, G., and Thiessen, D. S. Behavioral genetics: Method and research. New York: Appleton-Century-Crofts, 1969.)

Nichols, R. C., and Bilbro, W. C. The diagnosis of twin zygosity. Acta Genetica et Statistica Medica, 1966, 16, 265-275.

Tuddenham, R. D. Soldier intelligence in World Wars I and II. American Psychologist, 1948, 3, 54-56.

Vandenberg, S. G. Assortative mating, or who marries whom? Behavior Genetics, 1972, 2, 127-137.

Waller, J. H. Differential reproduction: Its relation to IQ test score, education, and occupation. Social Biology, 1971, 18, 122-136.

Zajonc, R. B. Family configuration and intelligence. Science, 1976, 192, 227-236.

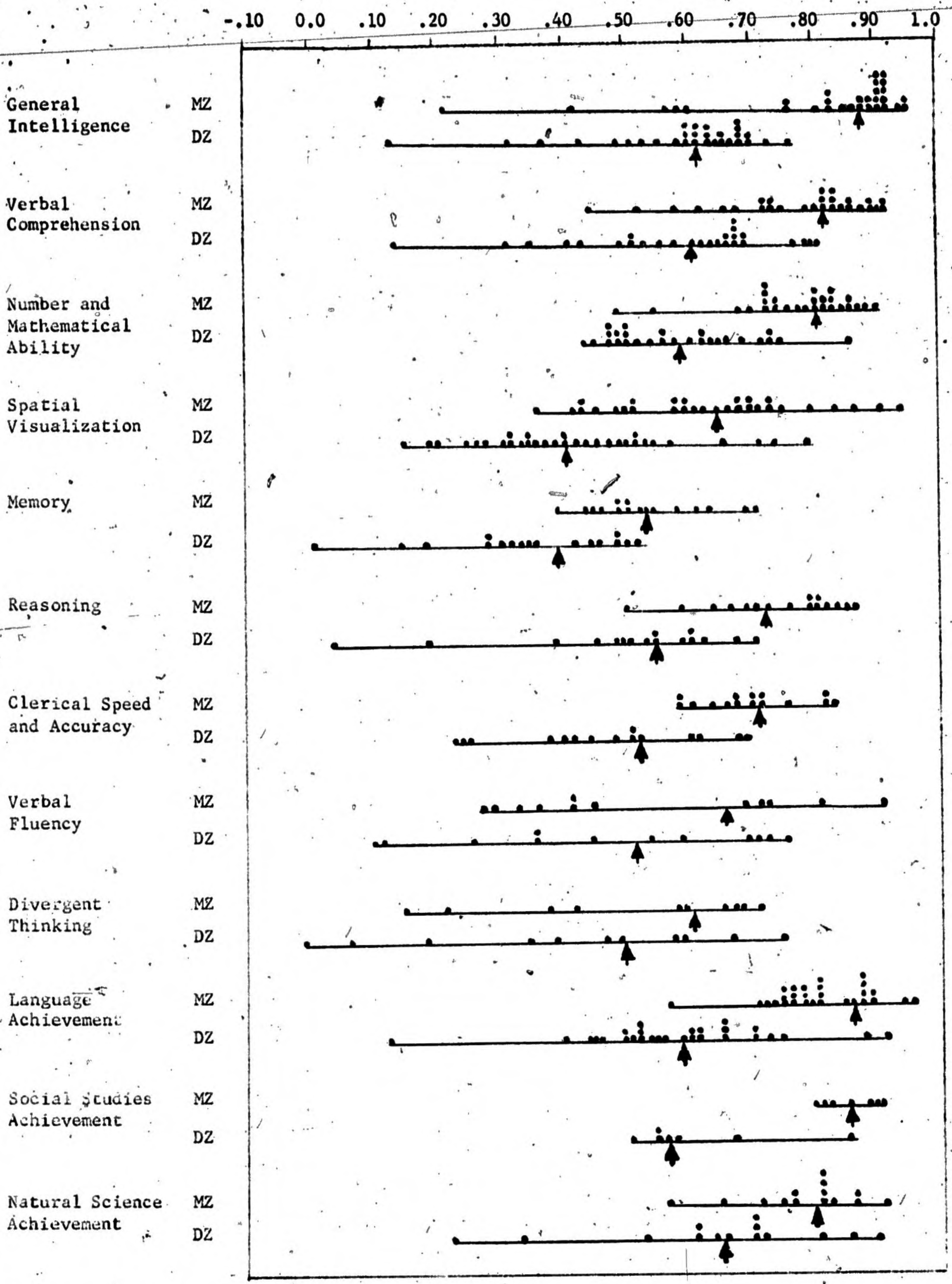

Figure 1. Intraclass correlations from twin studies of various abilities. Correlations obtained in each study for MZ and DZ twins are indicated by dots; the mean correlation, weighted by the number of cases, is indicated by an arrow below the horizontal line representing the range of correlations for each trait.

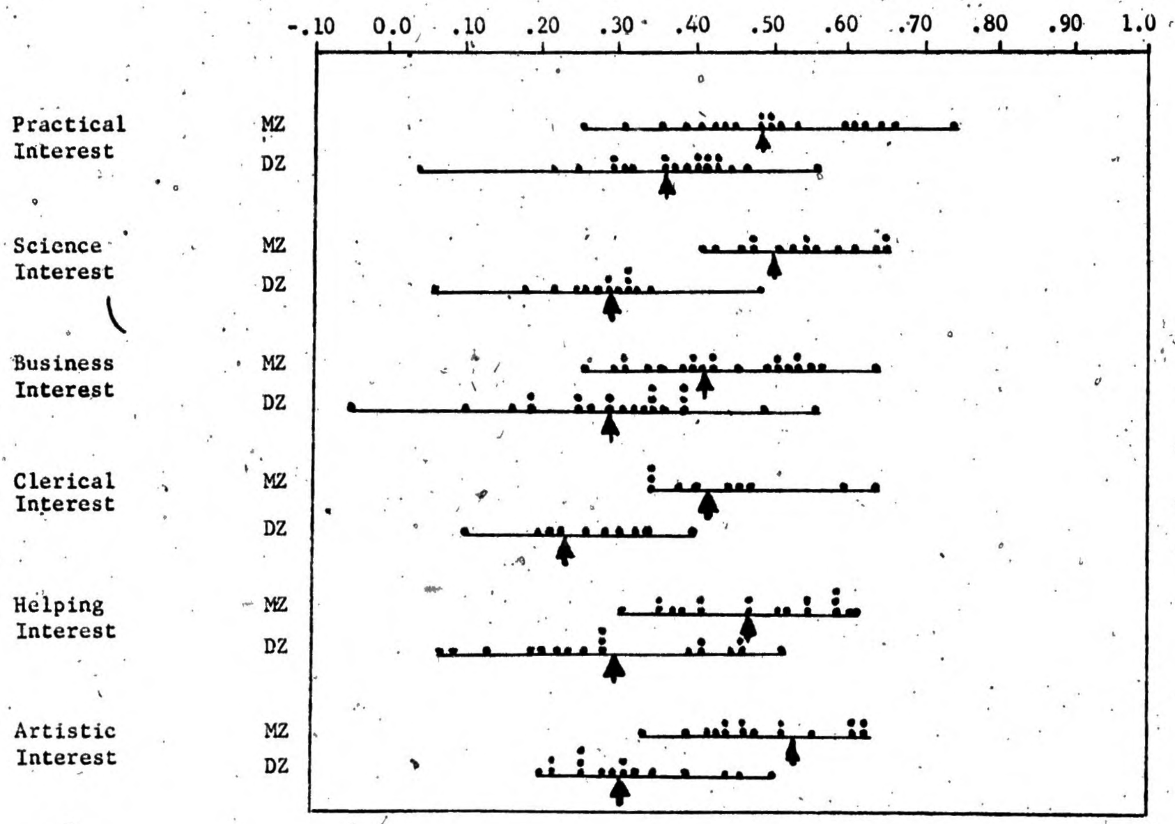

Figure 2. Intraclass correlations from twin studies of various interests. Correlations obtained in each study for MZ and DZ twins are indicated by dots; the mean correlation, weighted by the number of cases, is indicated by an arrow below the horizontal line representing the range of correlations for each trait.

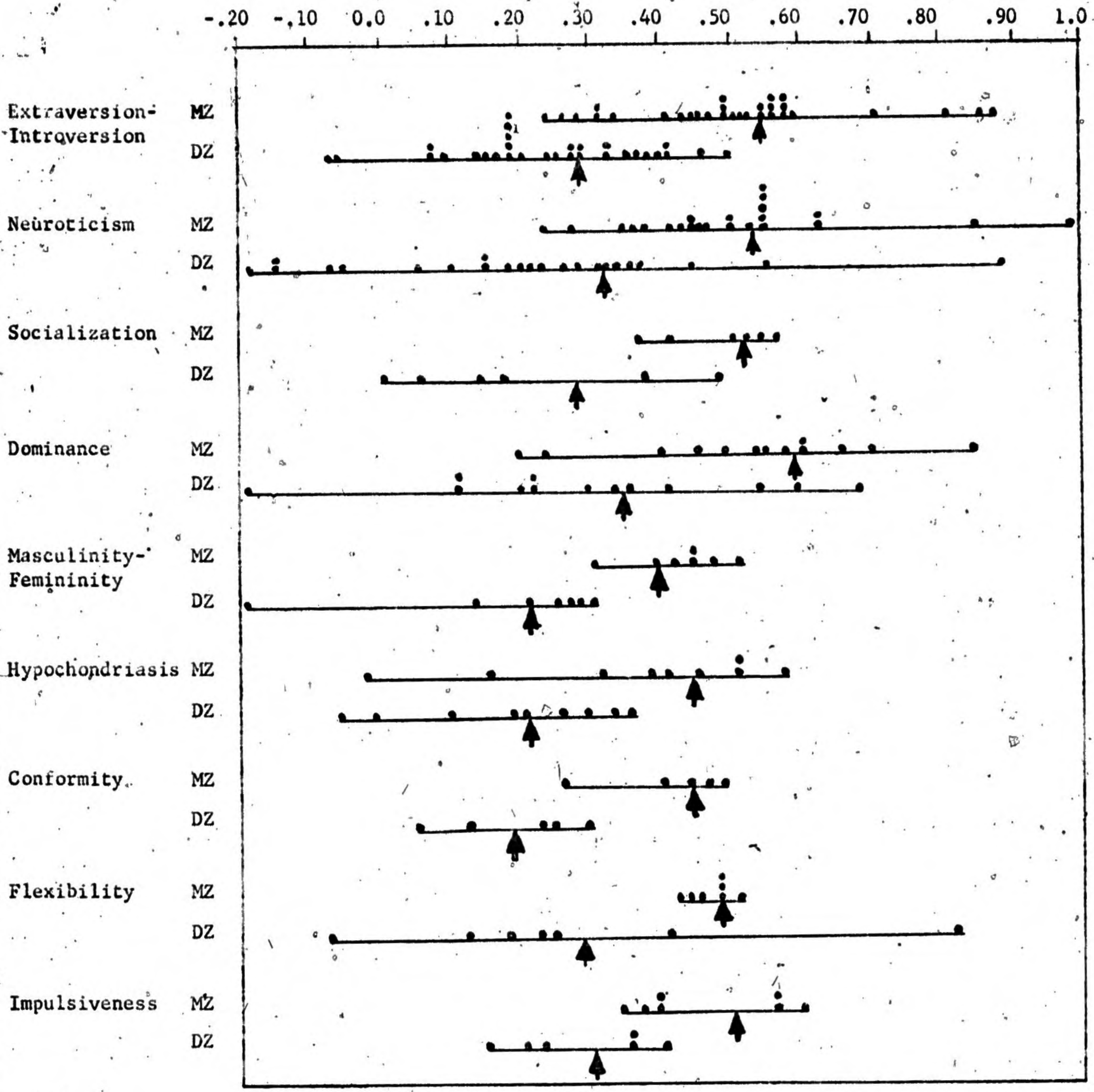

Figure 3. Intraclass correlations from twin studies of various personality dimensions. Correlations obtained in each study for MZ and DZ twins are indicated by dots; the mean correlation, weighted by the number of cases, is indicated by an arrow below the horizontal line representing the range of correlations for each trait.

Table 1

Mean Intraclass Correlations from Twin Studies of Various Traits

Trait	Number of studies	Mean Intraclass Cor. r_{MZ}	Mean Intraclass Cor. r_{DZ}	Difference $r_{MZ}-r_{DZ}$ Mean	Difference $r_{MZ}-r_{DZ}$ Stand. Dev.
Ability					
General Intelligence	30	.82	.59	.22	.10
Verbal Comprehension	27	.78	.59	.19	.14
Number and Mathematics	27	.78	.59	.19	.12
Spacial Visualization	31	.65	.41	.23	.16
Memory	16	.52	.36	.16	.16
Reasoning	16	.74	.50	.24	.17
Clerical Speed and Acc.	15	.70	.47	.22	.15
Verbal Fluency	12	.67	.52	.15	.14
Divergent Thinking	10	.61	.50	.11	.15
Language Achievement	28	.81	.58	.23	.11
Social Studies Achivement	7	.85	.61	.24	.10
Natural Science Ach.	14	.79	.64	.15	.13
All abilities	211	.74	.54	.21	.14
Interests					
Practical Interest	20	.50	.37	.13	.15
Science Interest	15	.54	.29	.25	.11
Business Interest	22	.45	.30	.15	.14
Clerical Interest	10	.44	.26	.18	.09
Helping Interest	18	.48	.30	.18	.14
Artistic Interest	16	.50	.32	.18	.13
All interests	116	.48	.30	.18	.13
Personality					
Extraversion-Introversion	30	.52	.25	.27	.14
Neuroticism	23	.51	.22	.29	.21
Socialization	6	.49	.23	.27	.17
Dominance	13	.53	.31	.23	.18
Masculinity-Femininity	7	.43	.17	.27	.21
Hypochondriasis	9	.37	.19	.18	.28
Conformity	5	.41	.20	.22	.15
Flexibility	7	.46	.27	.19	.27
Impulsiveness	6	.48	.29	.19	.12
All personality	106	.48	.29	.19	.12

Note: Mean correlations are unweighted averages of the studies involved. Because most twin studies employ multiple measures, the same twin sample may be represented in several traits.

Table 2

Median Intraclass Correlations for Various Trait Domains

Trait	Intraclass Correlation		
	Identical	Fraternal	Difference (I-F)
General Ability (NMSQT total score)[a]	.86	.62	.24
Special Abilities (5 NMSQT subtests)	.74	.52	.22
Activities (17 activities clusters)	.64	.49	.15
Interests (88 strong scales, male) (69 strong scales, female)	.53	.27	.26
Personality (27 CPI scales)	.50	.28	.22
Goals and Ideals (31 clusters of life goal, ideal-self and interest items)	.37	.20	.17
Self Concept (15 clusters of self concept ratings)	.34	.10	.22

[a]NMSQT is National Merit Scholarship Qualifying Test.

Table 3

Median Intraclass Correlations for Various Trait Domains
Corrected for Unreliability of Measurement and Errors of Diagnosis

Trait	Reliability of Measurement[a]	Intraclass Correlation: Identical	Fraternal	Difference (I-F)
General Ability (NMSQT total score)	.95	.92	.63	.29
Special Abilities (5 NMSQT subtests)	.88	.86	.57	.29
Activities (17 activities clusters)	.70	.93	.68	.25
Interests (88 strong scales, male) (69 strong scales, female)	.85	.64	.29	.35
Personality (27 CPI scales)	.80	.65	.33	.32
Goals and Ideals (31 clusters of life goal, ideal-self and interest items)	.65	.59	.29	.30
Self Concept (15 clusters of self concept ratings)	.65	.55	.13	.42

[a] Median test-retest reliabilities estimated from data provided in the test Manual, when available, or from internal consistency reliability calculations. These values were used in correcting the intraclass correlations for attenuation.

Table 4

Correlation of Twin Environments Implied by the Corrected Intraclass Correlations Shown in Table 3

Trait	Environmental Correlation: Identical	Fraternal	Average (I+F)/2
General Ability (NMSQT total score)	.73	.77	.75
Special Abilities (5 NMSQT subtests)	.65	.68	.66
Activities (17 activities clusters)	.83	.95	.89
Interests (88 strong scales, male) (69 strong scales, female)	.10	-.02	.06
Personality (27 CPI scales)	.13	.08	.10
Goals and Ideals (31 clusters of life goal, ideal-self and interest items)	-.02	-.02	-.02
Self Concept (15 clusters of self concept ratings)	-.12	-.42	-.27

Note: Environmental correlations were calculated from the corrected intraclass correlations in Table 3 using the formula in footnote 7. The calculation for general ability assumed a heritability of .70 and a genetic correlation for fraternal twins of .57. The calculation for all other traits assumed a heritability of .60 and a genetic correlation for fraternal twins of .50.

Table 5

Children Ever Born to Women 35 to 44 Years Old
by Education of the Woman: 1960 and 1974

Years of school completed		Number of women, 1974 (thousands)	Children ever born per 1000 total women (including single & childless) 1960	1974	Percentage change 1960-1974
Elementary:	Less than 8 yrs	697	3,205	3,778	17.9
	8 years	658	2,717	3,622	33.3
High School:	1 to 3 years	2,038	2,548	3,460	35.8
	4 years	5,517	2,243	2,834	26.3
College:	1 to 3 years	1,357	2,207	2,748	24.5
	4 years or more	1,346	1,918	2,163	12.8

Source: U. S. Bureau of the Census, Current Population Reports, series P-20, Nos. 263 and 277.