Record: 1

Title: Biological factors and psychometric intelligence: A review.

Authors: Johnson, Fred W.

Source: Genetic, Social & General Psychology Monographs. Aug91, Vol. 117 Issue 3, p315. 43p. 12 Charts.

Document Type: Article

Subjects: PSYCHOMETRICS
BODY size
INTELLECT

Abstract: Examines the correlations between psychometric intelligence and head and body size. Evidence of pleiotropy in the correlations; Physical correlates of psychometric intelligence; Within-family analysis of intelligence and head size.

Lexile: 1430

Full Text Word Count: 17937

ISSN: 8756-7547

Accession Number: 9607211186

Database: MasterFILE Premier

BIOLOGICAL FACTORS AND PSYCHOMETRIC INTELLIGENCE: A REVIEW

ABSTRACT. Results from a synthesis of correlations between psychometric intelligence and two physical traits, head size and body size, are reported. Within-family studies are reviewed for evidence of pleiotropy, the effect of a single genetic factor on two traits. Studies are also reviewed to determine whether prenatal effects bias twin studies, leading to underestimates of genetic influence. An N-weighted mean partial correlation (controlling height) of. 10 between intelligence and head size was found. Using a method developed by Van Valen (1974), the correlation of intelligence and brain size was estimated as .29 based on all the intelligence/head-size studies of adults and adolescents, and .44 based on studies measuring intelligence with IQ tests. The N-weighted mean partial correlations (controlling age) of intelligence and height were. 18 for children and. 22 for adults. The within-family studies indicated that pleiotropy may contribute to the correlation of intelligence with head size and to the correlation of intelligence with body size. Prenatal effects are not an important source of bias in twin studies or for heritability estimates based on them.

TRAITS MAY BE CORRELATED as a result of either environmental or genetic influences. Environmental correlations occur when there are systematic differences in the social, cultural, or physical environment affecting the expression of a character. Such correlations may arise from a single environmental influence or from different environmental influences that are associated with a third environmental factor. For example, IQ and height might be correlated because a single environmental influence, nutritional differences, affects both cognitive development and physical development. But IQ and height might also be correlated through differences in nutrition and differences in intellectual stimulation that both covary with a third environmental factor such as social class. Similarly, the correlation between intelligence and head size could arise from the same or different environmental influences. Environmental correlations may occur in the absence of any genetic covaria-tion for the traits.

Genetic correlations can occur through two mechanisms: common assortment of genes and pleiotropy. Common assortment occurs when genes for different traits segregate together as a result of linkage, gametic phase disequilibrium, selection pressure, or cross-associative mating. Positive cross-assortive mating occurs when individuals above the population mean on one trait mate with persons above the mean on another trait. If in one generation, the tall mate with the bright, then height and intelligence will be positively correlated in the next generation. Negative cross-assortative mating occurs when individuals above the population mean in one trait mate with persons below the mean on the other trait, producing a negative correlation between the two traits. Correlations due to assortative mating persist only as long as the mating patterns giving rise to them are maintained and will change as mating patterns change.

Linkage is a form of common assortment that occurs when the loci influencing different traits are in close proximity on the same chromosome. Like correlations due to cross-assortative mating, correlations due to linkage are also impermanent because linkages are broken by the crossing-over process in meiosis, the probability of crossing over being inversely related to the proximity of the loci but eventually becoming complete in the total population. Linkages may occur within families, however, even if the loci are un-linked in the population.

Gametic phase disequilibrium occurs when two populations with different frequencies of alleles influencing two traits mix to form a single population. In the first generation following mixing, the allele combinations will be in disequilibrium in the population, and the two traits will be correlated. But this correlation is transient because after one generation of random mating, the disequilibrium giving rise to the correlation will vanish as the alleles are distributed in combinations reflecting their frequencies throughout the population.

Common assortment of genes may also result from selection for pheno-types that involve a complex of traits. For example, if tool making confers a selective advantage, then genes that contribute to the fine motor ability, hand-eye coordination, spatial visualization, and other skills involved in making tools may be assorted together. Correlations due to selection pressures will persist only as long as the traits confer a selective advantage or selection pressures are maintained.

The second mechanism producing genetic correlations, pleiotropy, occurs when a single gene affects two or more distinct phenotypic traits. The contribution of a pleiotropic locus to the correlation will depend upon how many other nonpleiotropic loci also contribute to the traits' covariance and the relative importance of the pleiotropic and nonpleiotropic genes to trait expression. Unlike correlations due to common assortment of genes, pleio-tropic correlations are permanent.

Although one or more types of common assortment may be more important in a particular family or population, pleiotropy is of great general importance because the traits are linked causally. Both traits covary in response to the same genetic factor. Correlations due to common assortment do not involve a shared causal factor; they merely indicate that the genes for two traits are passed together from one generation to the next. Because these genes are independent elements, they may be separated in succeeding generations either by changes in mating patterns or selection pressures or by crossing over.

Traits influenced by the same gene necessarily involve a common process. Jensen (1980b) noted that understanding how this process influences one trait may lead to an understanding of how it influences the other trait. For example, a positive pleiotropic IQ/head-size correlation, or the IQ/brain-size correlation underlying it, would indicate that IQ and brain size are causally related, that IQ increases with brain size or some aspect of it such as cortical surface area. A pleiotropic IQ/body-size correlation would indicate that a gene influencing physical growth also influences cognitive ability.

Note that pleiotropy operates differently in these two examples. In the brain size example, the pleiotropic gene influences intelligence as a consequence of its influence on brain size. This may be diagrammed as follows:

```
Gene(s)  [right arrow]  Organismic [right arrow]  Brain size [right
                        process
arrow] Intelligence
```

In the body size example, because height itself is unlikely to have an effect on intelligence, the pleiotropic gene probably influences the two traits separately as follows:

```
                                               Body size
                                           [arrow up]
Gene(s) [right arrow]  Organismic process
                                           [arrow down]
                                               Intelligence
```

The first diagram represents an indirect and the second a direct pleio-tropic effect on different phenotypic characters arising from the same gene-regulated process.

Because pleiotropic correlations are potentially more informative about underlying processes, it is important to be able to distinguish them from other correlations. Jensen (1980b) has suggested doing this using correlations within families. Such correlations are based on partitioning within-individual (wI) correlations into between-family (BF) and within-family (WF) components. The BF correlation is calculated with the summed scores of members of sibling or twin pairs, and the WF correlation is calculated with the signed differences between members of a pair. For example, the BF correlation of two traits for a sibling pair is computed from the sum of the scores of both siblings for Trait 1 and the summed scores of both siblings for Trait 2. The WF correlation is computed from the signed differences between the siblings on Trait 1 and the signed differences between siblings on Trait 2.

The pair sums and differences partition the covariance between two traits in the population into orthogonal components. The BF component reflects influences that make members of a pair similar to one another but different from the members of other pairs, and the WF component reflects influences that make members of a pair different from each other. Covariances between two traits based upon the sums and differences add to the total covariance based on deviations from the mean of each variable. Jensen (1980b) noted that the summed scores are more reliable and difference scores less reliable than single test scores and he provided formulas to correct for attenuation both the BF and WF correlations due to the differences in reliability. The correction changes the correlation only minimally, however, not materially affecting their interpretation, and has been done only by Jensen in one or two of his studies.

Pleiotropic traits may be identified from WF correlations for dizygotic (DZ) twins or full siblings. Pleiotropic traits produce WF correlations (as well as BF correlations) because with the same gene or genes influencing both traits, one cannot inherit the genes affecting one trait without inheriting the genes affecting the other trait. WF correlations will not arise from common assortment of genes due to assortive mating or selection because alleles at the different loci segregate randomly within families. Linkages within families would contribute little to the WF correlation of a polygenic trait in a large sample. Between-family environmental factors such as social class and cultural differences will produce correlations between but not within families because they are common to but do not differ among family members. Environmental factors producing differences between children within a family could result in a WF correlation, however. This would occur, for example, if an illness retarded both the physical and cognitive development of one sibling but not the other. Because of this, WF correlations are not proof of pleiotropy but are consistent with it, being a necessary but not sufficient condition to demonstrate it. If a WF correlation is zero or near zero, then pleiotropy plays no or only a small role in the association between the two traits in a sample or in a population.

Physical Correlates of Psychometric Intelligence

Methodological Issues

Reliability and validity of measures. There are a number of methodological issues to be considered when evaluating the studies in this review. Clearly, the physical and intelligence measures should be valid and reliable. For the earliest studies, done in the late 19th and early 20th centuries, psychometric tests were not available, and intelligence was measured with school grades or teacher ratings. Such measures have low reliability and are of questionable validity because they reflect achievement motivation and teacher perception as well as intelligence. Although the physical measures would be expected to present fewer measurement problems because they are valid, in the case of external head size, obtaining reliable measurement is surprisingly difficult. Head circumference may at first glance appear to be the best measure of external head size, but because of interference from hair, especially in girls, and differences in technique, it may be less reliable than head length or breadth obtained with calipers.

Another problem with external head measures is their low correlation with brain size due to variation in skull thickness and the membranes surrounding the brain. In three studies cited by Van Valen (1974), the correlation between external head measures and brain size was only .48, a low correlation when one measure (external head size) serves as an index of another (brain size). Because the correlation of brain size with external cranial size is low, the intelligence/brain-size correlation, the correlation of interest, will be seriously underestimated when obtained using external cranial size as a proxy for brain size. As will be seen later, Van Valen has proposed a method for estimating the correlation of intelligence with brain size, taking the low head-size/brain-size correlation into account.

Factors affecting correlations. Three factors affecting the magnitude of correlations must also be considered when evaluating the studies in the following review: age, restriction of range, and combining groups. Especially in studies of children, age differences must be controlled. If physical variables were in fact unrelated to intelligence, they would nevertheless appear correlated with an unstandardized intelligence measure in a sample of children of varying ages because the older, larger children would obtain higher scores. Age must therefore be controlled either by adjusting the raw data for age differences, for example, with regression; by selecting samples from a narrow age range; or by computing partial correlations. In studies of the intelligence/head-size correlation, body size must also be controlled. This is most easily achieved by partialling height and weight from the zero-order correlation of head size and intelligence.

In samples selected for high ability, such as college students, the true correlation of intelligence and physical measures in the population will be underestimated. This is because the range of variation in selected samples is restricted. To obtain the true correlation of intelligence with physical traits, the full range of ability must be sampled. One indicator of the ability distribution in the sample, when the raw data are not available, is the standard deviation of the intelligence measure. If it approaches the normed standard deviation of the test, then it is likely that there is an adequate distribution of ability in the sample.

Combining groups such as men and women or Blacks and Whites may either increase or decrease correlations, depending on the relationship between the variables in the groups and on whether there are group differences on the measures correlated. For this reason, groups should be analyzed separately or group differences controlled statistically with regression. If groups are analyzed separately, the weighted mean correlation for all the groups can be calculated to obtain a result for the total sample.

Research Synthesis

For the correlations reported under each of the topics reviewed, I performed a research synthesis to summarize the results of the different studies by obtaining the average of the reported correlations, weighting each by its sample size. I have computed N-weighted means both for zero-order correlations and first-order partial correlations, the latter calculated with the standard formula presented by Marascuilo and Levin (1983), and I have presented the 95% confidence intervals for the N-weighted mean zero-order and partial correlations and the 95% confidence intervals for the correlations from each study. (Where a directional hypothesis is tested, as in the within-family studies, I gave 97.5%, one-tailed confidence intervals.) This method of research synthesis, averaging reported correlations, differs from a formal meta-analysis by not including an analysis of the distributional characteristics of the results, such as the variance among the reported correlations.

When the correlation between traits in the population is low, the results in any one study may not be statistically significant, and it may be concluded that the variables are not related at all. If a number of studies have been done, however, and the correlations are predominantly on one side of zero rather than varying randomly about it, the distribution of the correlations may provide evidence of a real relationship between the variables in the population. If a research synthesis yields a mean correlation that, although low, is statistically significant by virtue of its experimental replications, then it may be concluded that the variables are correlated. The weighted mean of the correlations will provide the best available estimate of the direction and magnitude of the correlation in the population.

This is the case with intelligence and physical variables. The two-tailed confidence intervals in the tables herein are 95%, the one-tailed limits are 97.5%. These confidence limits give both the one- and two-tailed confidence limits the same critical value, making the one-tailed limits and the upper two-tailed limits equal.

The following review and research synthesis is based on a thorough search of the literature, although it is not claimed to have been exhaustive. Note that, of the studies found, none reporting correlations were excluded from the review or research syntheses.

Intelligence and Head Size

The belief that the intelligence of individuals is related to the size of their head is a long-held one. It was given scientific support in the 19th century by Franz Joseph Gall, who originated the still currently accepted doctrine of localization of function but also founded the pseudoscience of phrenology. The belief was perhaps most widely held during the late 19th and early 20th centuries but waned in subsequent decades under attack by social scientists. The belief may have gained credence from the observation of 19th-century paleontologists that the farther back in evolutionary time they traced the lineage of mammalian species, the smaller were the species' cranial volumes relative to their head and body sizes.

According to Jerison (1982), "Encephalization is a measure of processing capacity after an adjustment for body size. It is, thus a measure of intelligence..." (p. 767). Hence, the larger the brain size relative to the body size typical of a species, the higher the species' intelligence. Jerison noted that the strong relationship between processing capacity and brain size found between species exists only weakly within species and predicted that the correlation between psychometric intelligence and brain size in humans would be quite low, .01 to .05. A review of the literature on humans shows that, when adjusted for body size, intelligence and brain size indeed appear correlated more highly than Jerison predicted.

The earliest quantitative studies of the intelligence/head-size relationship were done in the late 19th and early 20th centuries and were first reviewed by Whipple (1914). These and later studies were included in a review and critique by Paterson (1930) of the entire area of research correlating intelligence and physical attributes.

Paterson cited two methodological shortcomings common to many of the early studies: (a) reporting group means rather than correlation coefficients or information on within-group variability and (b) using age-grade location or teacher ratings rather than intelligence tests to measure mental ability. Certainly, these were necessarily shortcomings of the earliest studies because the correlation coefficient and intelligence tests had not yet been developed. Paterson described a number of studies, including those of Gaiton in 1888, the first quantitative study of the intelligence/head-size correlation, and of Porter in 1892, the first American study, showing that intelligence and head size are positively related. Although these studies used only the crude methods available to the earliest researchers, Paterson himself conceded that the consistency with which the studies found higher grades and teacher ratings associated with greater head size indicated that a real though low relationship exists between intelligence and head size.

Research Synthesis

Table 1 shows the results of the studies that reported correlations. Studies of young children appear separately in the table from those of adolescents and adults because the expression of these traits is least stable in the very young. But with the exception of one study (Weinberg, Dietz, Penick, & McAlister, 1974), the correlations for the very young do not differ appreciably from those for older subjects. The table shows that age was controlled in all the studies and sex in all but two. Although height was controlled in only a few studies, I computed ric.H, the partial correlation of intelligence (I) with external cranial size (C) controlling height (H) for each study using the standard formula (Marascuilo & Levin, 1983), which requires r_{IC}, r_{IH}, and r_{CH}. This was done with the value for r_{IC} from the respective study and the value for r_{IH}. from the respective study if available or if not, with the N-weighted mean correlations of. 18 and .22, respectively, for children and adults from Tables 4 and 5. The value for r_{CH} was obtained from the N-weighted mean, .31, of the correlations reported in Lee, Lewenz, and Pearson (1902), Reid and Mulligan (1923), and Passingham (1979).

All the correlations in Table 1 are positive despite methodological inadequacies lowering correlations in many of the studies, particularly the early ones, in which selected samples and unreliable intelligence measures were used. Three of the zero-order and five of the partial correlations are nonsig-nificant, but this is expected for single studies of a low correlation. With the exception of Weinberg et al. (1974), the correlations are similar whether the samples were selected or unselected, ethnically homogeneous or ethnically mixed, of children or adults.

For all the studies of adolescents and adults except Jensen's (1987), in which he computed a multiple R, the N-weighted mean zero-order and partial correlations (controlling height) were . 14(-+.02) and . 10(-+ .02), respectively. For the five adolescent and adult studies measuring intelligence with IQ tests, the weighted mean zero-order and partial correlations were higher, .21(-+.03) and. 17(-+ .03), respectively. The highest correlations (excluding Jensen's multiple R) were found in the two methodologically soundest studies, Murdoch and Sullivan (1923) and Suzanne (1979).

Murdoch and Sullivan's 595 subjects were 1 st through 12th grade pupils of middle- and upper-middle-class backgrounds. Children of higher socioeco-nomic (SES) origins were overrepresented in this sample, but the standard deviation of IQ, 16.5, was similar to the tests' normed standard deviation of 15, so there was no restriction of range in the distribution of intelligence. Standardized intelligence tests were used, and age and sex differences were controlled using within-group deviation scores. The zero-order and partial correlations obtained of intelligence with head size were .22 and. 18, respectively.

Suzanne's sample of 2,071 subjects was randomly selected from the 43,452 Belgian men aged 17 to 25 who were liable for military service in 1963 and was fully representative of the socioeconomic and geographic distribution of that country's male population. Intelligence was measured with the normalized total score of five subtests, and height was controlled. The zero-order correlations (partial correlations controlling height in parentheses) of the IQ measure with three head-size measures, circumference, length, and width, were .24 (.19), .22 (.18), and .13 (. 11), respectively. The median zero-order and partial correlations, those for head length, are the same as the zero-order and partial correlations Murdoch and Sullivan found.

None of the investigators in the adolescent or adult studies examined the linearity of the intelligence/head-size relationship. But in their analysis of data on large samples of 4-year-olds (approximately 10,000 Whites and 12,000 Blacks) from the National Collaborative Perinatal Project (NCPP), Broman, Nichols, and Kennedy (1975) found substantial nonlinearity. Nonlinear relationships are incorrectly estimated by the Pearson correlation. If the relationship is also nonlinear among older subjects, it could be either over-or underestimated by the correlations shown in Table 1 and by the N-weighted mean correlations based on them, depending on the nature of the nonlinearity.

Head Size and Socioeconomic Status

A correlation between intelligence and head size could arise from social class differences in nutrition or health care or perhaps from correlated differences in nutrition and intellect stimulation. The Broman et al. (1975) study permits a rough test of these environmental hypotheses by comparing correlations both within and across SES levels. Table 2 shows, for Blacks and Whites, the correlation between Stanford-Binet IQ scores and head circumference in the total group and in the lowest quartile, middle two quartiles, and highest quartile of an SES index based on family income and the education and occupation of the head of household. The correlations within social class levels are only slightly lower than the correlations in the total groups across SES levels. Although such a course grading of SES does not control for all social class variables, if the correlation were due to socially patterned influences such as nutrition and health care, it should be considerably lower within social class levels. That it is not suggests that the correlation is more influenced by genetic than. environmental factors. Note that although it might be tempting to partial SES from the IQ/head-size correlation, doing so would in large part remove IQ itself from the correlation because IQ is so highly correlated with SES, .50 to .70 in a number of studies surveyed by Jensen (1980a).

Head Size, Brain Size, and Intelligence

These researchers investigated the correlation of intelligence with head size only because the variable of interest, brain size, could not be measured directly. Van Valen (1974) provided mathematical proof of a formula to estimate the correlation of intelligence with brain size based on the assumption that head size is correlated with intelligence only through the relationship of head size and brain size. Letting I denote intelligence, B brain size, and C external cranial size, he proved that Rho_{IC}=Iain X Rho_{BC}, from which it follows that Rho_{IB} = Rho_{ic}/Rho[Sub BC]. Using a mean value for P_{BC} of .48, based on three studies, and a value for Pic of. 10, based on the six earliest studies listed in Table 1, Van Valen obtained an estimate for Rho_{IB} of. 10/.48 = .20. Correcting this value for attenuation to remove the effects of measurement error, which would be great because of the poor measures of intelligence used in most of the early studies, he arrived at. 30 as the best estimate of the correlation between intelligence and brain size.

Applying Van Valen's formula to the studies reviewed here (but not correcting for attenuation) yields estimates for P_{IB} of .14/.48 = .29 based on all the studies of adults and adolescents, and of. 21/.48 = .44 for the five studies that measured intelligence with psychometric tests. Note that these calculations have been done with zero-order correlations. Performing them with partial correlations controlling body size as Jensen and Sinha (1991) suggested would yield a lower estimate of P_{IB}, but Van Valen's proof may not extend to partial correlations (L. A. Marascuilo, personal communication; June 6, 1989). Because body size was not controlled, the values of .29 and .44 are likely to be overestimates.

In one study, brain size has been measured directly in living subjects. Using magnetic resonance imaging (MRI), Willerman, Schultz, Rutledge, and Bigler (in press) compared two groups of college students at the extremes of the intelligence distribution, one with low Scholastic Aptitude Test (SAT) scores (-<940) and Wechsler IQs (-< 103) and one with high SATs (> 1350) and Wechsler IQs (-> 130). In the total sample of 40 subjects, the correlation of brain size and intelligence was .51, and .35 when corrected for selection of extreme groups. The corrected correlation accords well with Van Valen's estimate of .30 and with the estimates of .29 and .44, all obtained indirectly from the correlation of intelligence with head size.

Using direct measures, the Willerman et al. (in press) study corroborates the results of studies that have only measured brain size indirectly and rather poorly as indicated by the brain-size/external-cranial-size correlation of .48 cited by Van Valen. The similarity of the Wilierman et al. results to the Van Valen estimate also supports Van Valen's assumption that head size is correlated with IQ only through the relationship of head size with brain size. Squaring the two more conservative estimates (.29 and. 30) of the IQ/brain-size correlation to obtain the coefficient of determination suggests that, in the populations sampled, about 10% of the variance in intelligence is due to individual differences in brain size or to a feature of the brain, such as cortical surface area, directly related to brain size.

Within-Family Analyses of Intelligence and Head Size

This review and the Willerman et al. (in press) report show that intelligence is positively correlated with head size and brain size within individuals. Although this finding is interesting, it takes on much greater significance if the correlation also occurs within families at a similar magnitude as within individuals. As explained earlier, this would be consistent with pleiotropy, suggesting that the traits share a common genetic determinant. Alternatively, a small or zero correlation within families would suggest that pleiotropy contributes little or not at all to the relationship of IQ within head and brain size.

Table 3 shows the results of the within-family analyses I used for this review, two of like-sexed DZ twins and one of full siblings. The table shows both WF correlations and, for comparison, the corresponding WI correlations (in which all pair members are treated as singletons; N = 2NP, NP = number of pairs), which the WF correlations would not be expected to exceed.

The Johnson (1990a) study used NCPP twin data (the data used by Broman et al., 1975), in which the children's head circumference and cognitive ability were measured at ages 4 and 7. Jensen's (1987) results were based on data for 61 same-sexed DZ pairs and 82 MZ pairs from the Louisville Twin Study. He obtained the multiple correlation between three measures of head size (length, breadth, and circumference) and scores from the unrotated first principal component of 16 cognitive tests. Johnson's (1990b) sibling analysis was based on two cohorts of children, one tested at age 5 on the Colored Progressive Matrices (CPM) and Peabody Picture Vocabulary Test (PPVT) and the other at age 10 on the Standard Progressive Matrices (SPM) and PPVT by the Child Health and Development Studies, a longitudinal public health study. Age, race, and sex were controlled in all the correlations in the table, but height was controlled in none of them.

The correlations were not corrected for attenuation due to the lower reliability of difference scores, but correction would change the correlations little and not materially affect their interpretation. Note that in the Jensen (1987) study, the WI correlations were computed with the total twin sample in the study including both MZ and DZ twins, not just the DZ twins used in the WF analyses. This should not affect the magnitude of the

WI correlations, however, because there is no evidence that the IQ/head-size correlation differs between MZ and DZ twins.

Table 3 shows that for the studies reporting bivariate correlations, the mean WF correlation is much smaller (.06) than the mean WI correlation (. 19) and is nonsignificant. Obtaining the ratio r_{WF}/r_{WI} indicates the proportion of the correlation within individuals (r_{WI}) that also occurs within families (r_{WF}). For the weighted mean bivariate correlations, this ratio is r_{WF}/r_{WI} = .06/. 19 = .32, or 32%, indicating that about one quarter of the WI correlation occurs within families. Not included in the mean WF bivariate correlations is Jensen's (1987) WF multiple correlation of. 28 (WI multiple R,. 31), which is nearly significant and has a R_{WF}/R_{WI} ratio of 93%. Taken together, the weighted mean WF bivariate correlation of .06 and WF multiple correlations of. 28 suggest that pleiotropy may contribute to the correlation of intelligence with head size.

Intelligence and Body Size

Although body size, unlike head size, may not be intuitively associated with intelligence, scientific writers in the late 19th century lent their authority to the notion that mental and physical superiority go hand in hand. Galton, in particular, emphasized the superior physical attributes of the eminent men he studied. The first quantitative studies of the relationship between intelligence and body size were done not long after Galton's book Hereditary Genius appeared in 1869. Whipple (1914) and Paterson (1930), who did the earliest reviews of the head-size studies, were also the first to review the studies of the relationship between intelligence and body size. Paterson noted that the early studies of intelligence and body size were similar methodologically to the early head-size studies and shared with them the two chief shortcomings described earlier: (a) reporting group means rather than correlation coefficients and (b) measuring intelligence with age-grade location or teacher ratings rather than IQ tests. As with the head-size studies, neither the correlation coefficient nor intelligence tests had been invented at the time of the earliest body size studies. In general, despite the crude methodologies available to researchers, the early studies showed that brighter children are taller and heavier.

Research Synthesis

The results of the studies that reported correlations of intelligence and height are summarized for children in Table 4 and for adults in Table 5. The results for children and adults are presented separately because the correlation of intelligence with body size in children may be attenuated by spurts and lags in both physical and cognitive development. In the two tables, there are only a few nonsignificant correlations and one that is negative, all in studies with relatively small or highly selected samples. Most of the studies of children controlled for the inflating effects of age through partial correlation or other means, but only a few of the adult studies did so. However, in all the recent (post-1950) adult studies except Passingham (1979), the subjects were military recruits from a narrow age range and near their maximum adult height. Removing the effects of age from these correlations would change them minimally if at all.

Note that although age-adjusted correlations should be lower because one source of covariance is removed, r_{IH}.A (' 18) is greater than r_{IH} (. 14) in Table 4 because these two mean correlations are based on different studies, with the value for r_{IH} determined primarily by the Wilson et al. (1986) study, which had a large sample. Sex was controlled in all the adult studies primarily by selecting all-male samples, and in all the children's studies except Broman et al. (1975). Most of the samples were unselected, but with the exception of Lee et al. (1902), positive correlations were also found in all the samples selected for mental ability. With a few exceptions, similar correlations of intelligence with body size were found in samples drawn both from populations selected and unselected for mental ability, in ethnically homogeneous northern European samples, and in ethnically and racially mixed American samples. The N-weighted mean partial correlation for children, .

18, was lower than that for adults, .22, a significant difference ($p < .001$) that may reflect developmental instabilities during childhood.

Fewer studies reported correlations of intelligence with weight than with height. Table 6 summarizes the results for the weight studies of both children and adults. Age was controlled in most of these studies and sex in all but Broman et al. (1975). Of the two zero-order correlations, one is negative (though nonsignificant and based on a small N); and of the nine partial correlations, four are nonsignificant. In general, the correlations are lower than those of intelligence with height. This may be because weight is related to intelligence through its relationship with height but has a lower correlation because of variation in body fat unrelated to stature. The N-weighted mean correlation for the studies of children, excluding Broman et al. (1975), who combined males and females, is . 14 (-+ .02). There are too few studies of adults to obtain their weighted average.

Several of the studies of children were both methodologically sound (used standardized tests and unselected samples; controlled age and sex) and had quite large samples. The Scottish Mental Survey is probably the best study because of the size and representativeness of its sample. The Scottish Council for Research in Education (1953) administered the Moray House Test, a verbal reasoning, group intelligence measure, to all 11-year-old children in Scotland in 1932 and again to all l 1-year-old Scottish children in 1947. From the children tested in 1947, the 7,380 (3,674 boys and 3,706 girls) who were born on the first three days of each month were selected for more detailed study including measurement of their height and weight. Comparisons on all the measured variables were done between the sample of 7,380 (the lower Ns in the tables are due to missing data) and the total population of 11-year-olds. As would be expected with such a large sample, no significant differences were found between the sample and the population. The correlations obtained were among the highest of those shown in Tables 3 and 4. The age-partialled correlations of test score with height were .24 and. 26 for boys and girls, respectively, and of test score with weight, . 16 and .22, for boys and girls, respectively.

The four largest adult studies were done in Scandinavian countries with fully representative samples of military conscripts who were tested on standardized measures. The sample in the Teasdale, Owen, and Sorenson (1990) study was particularly large, more than 71,000, and fully representative geographically and economically. Intelligence was measured with a four-part test whose total score correlates from. 77 to. 82 with the Danish WAIS (Wechsler Adult Intelligence Scale). The correlation of total score with height, controlling age, was .22, similar to the correlation for boys in the Scottish Mental Survey.

As with head circumference, there is evidence that the relationship between intelligence and body size is nonlinear. Broman et al. (1975) found a nonlinear regression of IQ on height and on weight in their large sample of 4-year-olds, and Teasdale et al. (1990) also found a nonlinear relationship in their very large sample of young adult males, a group at or near their maximum stature. These findings suggest that the Pearson correlation may underestimate or overestimate the relationship between intelligence and body size, depending on the nature of the nonlinearity.

The Coadvancement and Adult-Persistent Theories

In separate reviews, Tanner (1966, 1969) described two theories accounting for the correlation between intelligence and body size in children: the coad-vancement and adult-persistent theories. The coadvancement theory was originated by the anthropologist Franz Boas, who maintained that growth affecting both physical and mental development occurred at different rates in different children. Hence, larger, brighter children were accelerated for their age in both physical growth and intelligence, but smaller, duller children would catch up in both areas by adulthood. Porter maintained the adult-persistent theory, holding that those who were larger and brighter as children were also larger and brighter as adults and that the smaller, duller children do not catch up.

Douglas, Ross, and Simpson (1965), at Tanner's suggestion, directly tested the coadvancement hypothesis by analyzing the correlation between height and intelligence within levels of sexual maturation. If the correlation were due to a link between brain development and physical development, it should be zero or greatly diminished among children of the same sexual maturity. Douglas et al. obtained height and IQ data at three age levels (7 through 8, 11, and 15) on random samples of 1,475 boys and 1,389 girls. They grouped the children into three levels of sexual maturity and also four of SES and two of family size. The mean IQ-height correlation for the three age levels was. 14 and. 12 for the boys and girls, respectively. Similar correlations were found within levels of sexual maturity both within and across levels of the other three factors, sex, SES, and family size, indicating that the correlation is not due to coadvancement in intellectual and physical development but also occurs at fixed levels of sexual maturation.

The results in Tables 4 and 5 provide further evidence against the coad-vancement hypothesis. The hypothesis predicts that the intelligence/body-size correlation will be zero in adulthood, but the mean correlation for adults of .22 is not only nonzero, it is significantly higher than the correlation of. 18 for children as noted above. Douglas et al.'s (1965) finding that the IQ/body-size correlation occurs within maturation levels excludes genetic factors that control rates of development from a role in the correlation. Their finding also excludes environmental factors such as between-family differences in nutrition or health care from a role in the correlation because these differences would also be expected to affect rates of development. Their finding, however, is consistent with a genetic correlation due to positive assortative mating, which could give rise to a correlation across as well as within maturation levels.

Height, Intelligence, and SES

Just as intelligence and head size are correlated within SES levels, intelligence and body size are also correlated within social class levels. This has been shown in three studies. For example, Scott, Illsley, and Thomson (1956) obtained correlations within three social class levels in a study of 278 Aberdeen women experiencing their first pregnancy. The women were tested on both the Wechsler-Bellevue and Raven's Progressive Matrices. For the total group, the correlations of height with the Wechsler and Raven were .24 and .18, respectively. Within high, middle, and low social class levels, the correlations for the Wechsler were .04, . 16, and .25, respectively, and for the Raven -.08, . 14, and. 10, respectively, reduced, but not greatly, from the correlations in the total sample.

In Douglas et al's. (1965) study testing the coadvancement theory, they obtained correlations between IQ and height within four SES levels. These correlations were between .09 and. 15, and are nearly as large as the correlations across SES levels of. 14 and. 12 for boys and girls, respectively.

Broman et al. (1975) also obtained correlations of intelligence with body size measures both within and across social class levels. These correlations are shown in Table 7. As with the Broman et al. IQ/head-size correlations, the correlations of intelligence with height and weight are only slightly lower within SES levels than across SES levels. In all three of these studies, the correlations within SES levels are only slightly lower than the correlations across SES levels, suggesting that the IQ/body-size correlation is largely unaccounted for by factors presumed to vary between social classes, such as nutrition or intellectual stimulation. Like Douglas et al's. (1965) finding that height and IQ are correlated within levels of sexual maturation, this is consistent with a genetic correlation due to assortative mating.

Within-Family Analyses of Intelligence and Body Size

This review has shown that intelligence and body size are correlated over the entire course of development from early childhood through adulthood. Moreover, two lines of evidence suggest that the correlation may involve a genetic factor. First, the correlation is only slightly lower within SES levels, indicating that it involves factors other than environmental differences associated with social class. Second, the heritabilities of both traits are substantial, for height about .95 and IQ about .70. The covariance of traits whose variance in the population is largely genetic is more likely due to genetic than environmental influences. As noted in the introduction, a genetic correlation in the population between two polygenic traits such as height and IQ can arise from either common assortment of genes or pleiotropy, and if from the latter, the traits will be correlated within families.

For this review, I found seven within-family analyses of the IQ/body-size correlation, four of full siblings and three of same-sexed DZ twins. The twin studies were done with relatively small samples. Two of them (Jensen, 1987; Johnson, 1990a) were among the studies using within-family analysis of the IQ/head-size correlation. I performed the analysis for the fourth twin study on data reported by Thurstone, Thurstone, and Strandskov (1955). Scores from the unrotated first principal component of six PMA tests (Verbal, Number, Spatial, Word Fluency, Reasoning, and Memory) were used to measure intelligence in this analysis. Age, race, and sex differences on both the cognitive and physical variables were controlled using regressive procedures. The earliest sibling study (Laycock & Caylor, 1964) was done with 81 pairs in which one member was both gifted and had an IQ 20 points higher than the other member. Jensen's (1980b) study included large numbers of sibling pairs, both Black and White, of elementary school age, who were tested on seven measures.

There were also two racial groups in Nagoshi and Johnson's (1987) study: Americans of European ancestry (AEA) and Americans of Japanese ancestry (AJA). Intelligence was measured with scores from the unrotated first principal component of a battery of 15 diverse cognitive tests. In Johnson's (1990b) study, there were two samples of full siblings. One sample (NP = 310) was tested at three age levels: 5 years, 10 years, and during adolescence. Mean correlations for the three measurement occasions are shown in Table 8. The other sample (NP = 153) was tested once, at age 5.

Table 8 shows that the N-weighted mean of all the WF correlations is significant, though marginally. Laycock and Caylor's (1964) results and Jenson's (1987) WF multiple correlation are both nonsignificant, but that is to be expected in individual studies involving low correlations, as discussed earlier. The (marginally) significant mean WF bivariate correlation suggests that pleiotropy plays some role in the correlation of intelligence with body size. The r [subWF]/r_{WI} ratio indicates that about .04/. 13 = .31, or 31%, of the correlation within individuals also occurs within families.

Prenatal Biases in Twin Studies

Studying twins may be the most frequently used approach for estimating the relative contributions of heredity and environment to variation in human traits. The method is based on comparing the resemblance of monozygotic (MZ) and dizygotic (DZ) cotwins. Genetically, MZ cotwins are duplicates whereas DZ cotwins are no more alike than full siblings, sharing, on average, half their genes. In the absence of systematic differences in environmental variation, any greater resemblance of MZ over DZ cotwins must be due to their doubled genetic similarity. The validity of the assumption of equivalent environmental variances for the two zygosities is essential, for without it the inference of genetic causation of the greater MZ cotwin resemblance cannot be accurately or correctly made.

If the environmental variances for the MZ cotwins and the DZ cotwins are unequal, genetic influences may be either over- or underestimated, depending on how the variances differ. The equal environments assumption also applies to the prenatal environment. If intrauterine conditions are unequal for MZ pair members, this may result in lasting differences that reduce cotwin concordance and bias downward twin method estimates of genetic influence. Researchers have investigated and so far supported the soundness of the equal environments assumption with regard to the postnatal social environment (Loehlin & Nichols, 1976; Matheny, Wilson, & Dolan, 1976; Plomin, Willerman, & Loehlin, 1976; Scarr & Carter-Saltzman, 1979). The remainder of this article reviews research bearing on the soundness of this assumption with respect to the prenatal environment.

Prenatal Causes of MZ Cotwin Discordance

Price (1950) has suggested that a class of effects he termed primary biases operating in the prenatal environment may create persisting within-pair differences. The most important of these are the effects of the mutual circulation that occurs in monochorial MZ pairs. Price noted that because of the "chance factors of position and growth of placenta] vessels, as well as torsions in cord veins and arteries which, in the cords of nontwin embryos, are rarely of consequence," the circulation between cotwins is usually imbalanced. "As a re-suit, the development of either or both fetuses may be modified... and although the surviving twin infants recover from the condition to a large extent, it seems very probable that some of the effects are lasting" (p. 305).

Whether a monozygotic pair shares mutual circulation depends on when during its development the embryo splits to form twins. If splitting occurs early, during the first three days following ovulation and before the embryo implants in the uterus, the twins will be diamnionic and dichorionic, each twin surrounded by a separate amnion (inner membrane) and chorion (outer membrane). The placentas of the embryos will also be separate if the embryos implant far apart in the uterus but may become fused if they implant close together. If splitting occurs before the middle of the second week of gestation, after differentiation of the chorion but before that of the amnion, the twins will be monochorionic (sharing a single chorion) and diamnionic with a single placenta. Finally, if division occurs after the middle of the second week of gestation, when both the amnion and chorion have differentiated, the pair will be monochorionic and monoamnionic (sharing a single amnion), again with a single placenta. About 30% of MZ twins are dichorial. Of the 70% that are monochorial, only 3% are monoamnionic. In contrast, because they develop from separately fertilized ova, virtually all dizygotic twins are diamnionic and dichorionic. Their placentas may also be either fused or separate (Bulmer, 1970).

In the great majority of monochorial pairs (about 90%), the placental circulation of the twins is connected by vascular anastomoses, but only rarely is it connected in dichorial pairs. Mutual circulation is inefficient, and the vascular anastomoses frequently result in cotwin birth weight differences. One condition of mutual circulation, the twin transfusion syndrome, occurs when an artery from one fetus supplies the placenta of the other fetus and there is no compensating return circulation. At delivery, the transfusing, or "donor," twin is often anemic, and the transfused, or "recipient," twin plethoric (has excess blood) with the two twins usually, though not always, differing considerably in birth weight (Bulmer, 1970).

Several studies have found clinical evidence of the syndrome in about 30% of the monochorionic pairs samples (Strong & Corney, 1967). A high mortality rate is associated with the syndrome. Among the 19 cases diagnosed in a study of 130 monochorionic pairs by Rausen, Masako, and Strauss (1965), both twins died in 10 and 1 twin died in 5 pairs. In only 4 pairs, or 3% of the total monochorionic sample of 130 pairs, did both twins survive. Melnick, Myrianthopoulos, and Christian (1978) noted that because so few pairs survive the transfusion syndrome intact, it is unlikely that the syndrome by itself is an important source of intrapair differences.

The remaining question about the vascular anastomoses, then, is their effect on the development of the large proportion of monochorionic twins who share mutual circulation but are not diagnosed with the transfusion syndrome (James, 1982). Wilson (1979) had provided evidence that mutual circulation does affect fetal development. He found that MZ twins, although genetically alike, may differ greatly in birth weight and are less concordant than DZ twins in length and weight at birth.

As Wilson's (1979) findings suggest, birth weight is a useful index of the effects of prenatal conditions on fetal development. In singletons as well as twins, the effects of many prenatal conditions will be reflected in birth weight. Prenatal effects on the later development of traits of interest such as intelligence can be studied by comparing the IQ scores of individuals differing in birth weight. In the case of twins, prenatal effects may be investigated by comparing the mean scores of heavier and lighter birth weights. A second method of investigating prenatal effects on twins is to correlate intrapair differences on cognitive measures with intrapair differences on physical measures, such as height, under the assumption that both kinds of differences have a common prenatal origin. In this review, I have taken both approaches. An advantage of twins over nonsibling singletons is that variables such as gesta-tional age and between-family differences in prenatal care are controlled.

Chorion Type and Within-Pair Variability in Intelligence

In four of the five chorion type studies I found for this review, chorionicity was determined by the most reliable means, placental examination. Melnick et al. (1978) studied 53 monochorionic MZ pairs (23 White, 30 Black) and 33 dichorionic MZ pairs (9 White, 24 Black) whose intelligence was measured with seven subtests of the WISC (Weschsler Intelligence Scale for Children) at age 7 in the NCPP. Melnick et al. found that, among the White twins, the dichorionic pairs were significantly more discordant for IQ than the monochorionic pairs. There was a similar though nonsignificant trend in Blacks. Based on these findings, Melnick et al. suggested that the two types of MZ twins comprise separate populations with respect to IQ and that the greater within-pair variation of the dichorionic subpopulation may bias twin estimates of genetic influence.

Rose, Uchida, and Christian (1981) studied 17 monochorionic and 15 dichorionic adult MZ pairs, measuring their verbal and performance skills with the Vocabulary and Block Designs subtests of the WAIS. There was significantly greater intrapair variation for the dichorionic pairs than for the monochorionic pairs on Block Designs but not on Vocabulary subtests.

Welch, Black, and Christian (1978) obtained Bayley (1956) mental development scores at 18 months of age for 20 monochorionic and 12 dicho-rionic MZ pairs and found intrapair variation greater for the monochorionic twins but not significantly so. Welch et al. also described an unpublished study by Brown (1977) done with 55 monochorionic and 44 dichorionic MZ pairs tested on the Stanford-Binet at 4 years of age in the NCPP. Brown found intrapair variability unrelated to chorion type. It should be noted that the samples in the Brown and in the Melnick et al. (1978) studies must have overlapped, because the data for both studies were from NCPP twins, at age 4 in Brown and age 7 in Melnick et al. It is possible that Brown may have found intrapair variability related to chorion type had he performed separate analyses for Blacks and Whites as did Melnick et al. But it would also be expected that within-pair variation would be easier to detect at age 4 than at age 7 because twins recover from prenatal deficits as they mature, a phenomenon shown by Wilson (1979) and discussed later.

Breland (1974) studied MZ pairs who were tested on the five subscales of the National Merit Scholarship Qualifying Test (NMSQT) taken by college-bound high school juniors, diagnosing chorionicity on the basis of handedness. Her sample included 365 pairs (76%) concordant for handedness and therefore presumed dichorionic, and 117 (24%) discordant for handedness and presumed monochorionic. Using a multivariate test, she found no significant difference at the .01 level between vectors of mean within-pair differences for the dichorionic pairs and for the monochorionic pairs on the five NMSQT subtests. To account for the nonsignificant result, she suggested that handedness may be an inaccurate indicator of chorionicity and that pairs most discordant for IQ may have been excluded from this highly selected sample. Misdiagnosis of chorionicity also seems likely, because 70% of MZ twins are monochorionic but only 24% of her sample was presumed monochorionic based on the handedness criterion.

A finding related to these studies of intrapair variation was reported by Bailey and Horn (1986). They compared the variance of IQ scores for higher-and lower-scoring cotwins, reasoning that the disadvantageous prenatal effects unique to or more common in MZ twins, mutual circulation, for example, would have a greater effect on the lower-scoring cotwin. They used data from five MZ and three same-sexed DZ samples of adolescent and adult twins published in earlier studies, and found the variance of each test for the lower-scoring cotwins greater for the MZ but not DZ twins. This would indicate that there is extraneous within-pair variation in MZ twins, reducing their similarity relative to DZ twins and leading to underestimates of herita-bility.

Jensen (1987), however, was unable to replicate Bailey and Horn's (1986) results with the raw data from 22 MZ and 18 same-sexed DZ samples that included the samples from all studies in the twin literature publishing raw data, including those Bailey and Horn analyzed. Jensen concluded that Bailey and Horn's results must have been based on the atypical selection of studies from this population of published studies.

Table 9 summarizes the results of the five studies comparing the intrapair variability in intelligence of monochorionic and dichorionic MZ twins. In five of the seven comparisons, intrapair variability was greater in monochorionic pairs, but significantly so in only two of the comparisons. These results suggest that there may be a relationship between chorion type and intrapair variability, but it is weak and not an important source of MZ cotwin discordance potentially biasing twin studies. In addition, the results cast doubt on Melnick et al.'s (1978) suggestion that monochorionic and dichorionic twins form separate populations with respect to intelligence.

Birth Weight Differences and Intelligence

As noted earlier, birth weight is a useful index of prenatal conditions in twin pregnancies. Cotwin birth weight differences may be used as an index of prenatal conditions that have not been observed or recorded but the effects of which may be evident in cotwin birth weight discordance. More than 10 studies have compared the intelligence of lighter and heavier birth weight MZ cotwins to determine the degree to which prenatal effects are a source of bias in twin studies.

In an early study, Allen and Kallmann (1962) showed that low birth weight is related to severe mental deficiency. Comparing 12 MZ pairs discordant for institutionalization with 15 MZ pairs concordant for institutionaliza-tion, they found a significant birth weight difference between the discordant pairs but not the concordant pairs. This suggests that the etiology for mental defect differs in the two groups, with intrauterine conditions playing a major role in the discordant group and a handicapping gene, necessarily shared by both cotwins, playing a major role in the concordant group. The generaliza-bility of this study is limited, however, by the small number of pairs in each group and their status as mental defectives who are not part of the population of normal twins sampled for twin studies.

Babson, Kangas, Young, and Bramhall (1964) studied 9 MZ pairs (median age, 8.5 years) in which the smaller member weighed under 2,000 grams at birth and was at least 25% lighter than his mate. The lighter birth weight cotwins averaged 6.56 IQ points lower on the Stanford-Binet and 2.67 points lower on the PPVT

(Peabody Picture Vocabulary Test), both differences significant. Babson and Phillips (1973) followed up these twins during adolescence. At a mean age of 13 years, the lighter cotwins had WISC Verbal IQs and PPVT scores 8.7 and 7.4 points lower, respectively, than the heavier co-twins, both differences significant and greater than when initially tested.

Churchill (1965) studied 22 MZ pairs aged 5 to 15 recruited from a public school psychology clinic to which almost all had been referred for learning problems. This recruitment method produced a sample selected for low mental ability as is apparent in the mean IQs reported. Chorion type obtained from birth records and physical resemblance were used to diagnose the zygosity of 13 and 9 pairs, respectively. Mean Full Scale WISC IQs (in parentheses) of the lighter (80.9) and heavier (85.2) birth weight cotwins differed significantly ($p <. 0005$). The Full Scale IQ difference was largely due to the difference in mean Performance IQ (in parentheses) between the lighter (82.8) and heavier (89.0) birth weight cotwins, which was significant ($p < .0005$), and less to the difference in mean Verbal IQ (in parentheses) between the lighter (82.5) and heavier (84.4) birth weight cotwins, which was nonsignificant. In this study, no attempt was made to examine the relationship between intrapair birth weight and IQ differences, and the method of zygosity diagnosis for 9 of the pairs, physical similarity, was questionable.

Reexamining the data of Churchill's (1965) study, Willerman and Churchill (1967) noted that the relationship between IQ and birth weight was clearer when sets involving at least one breech delivery were excluded. They then repeated Churchill's study with only MZ pairs in which both members had normal vertex deliveries, 13 sets from the earlier study (mean age, 12 years) and 14 new sets recruited through a twin club (mean age, 9.5 years). In this combined sample, both the Performance IQs (in parentheses) of the lighter (89.1) and heavier (94.4) birth weight cotwins and the Verbal IQs (in parentheses) of the lighter (84.4) and heavier (88.8) birth weight cotwins differed significantly. Note that this study may have carried over from the Churchill study several of the pairs diagnosed only on the basis of physical similarity. Unlike Churchill, Willerman and Churchill did analyze the relationship between intrapair birth weight and IQ differences, and found their correlation, .02, to be an almost literally zero.

Scarr (1969) performed a study unusual for its pooling of samples and measure of intelligence. She combined the 27 MZ pairs of Willerman and Churchill's (1967) study (giving 13 pairs from Churchill's 1965 report, including any misdiagnosed, yet a third opportunity to contribute to science) with a new group of 25 female MZ pairs (mean age, 95 months). Scarr assessed the new twins' intelligence with the Draw-a-Person (DAP) Test, which yields a performance IQ correlating .40 to .75 with both the Verbal and Performance scales of the WISC, the test given the Willerman and Churchill twins. Thus, not only were different IQ measures used in the two groups comprising her total sample, the two tests had at most half their variance in common. Scarr found that the IQs of the heavier cotwins exceeded those of the lighter cotwins in a significantly disproportionate number of pairs. She also found that the magnitude of the cotwin IQ differences was positively related to the magnitude of the cotwin birth weight differences even though, as just noted, the differences were uncorrelated in the 27 Willerman and Churchill pairs that comprised more than half her total sample.

Kaelber and Pugh (1969a) began their paper by noting that in the Allen and Kallmann (1962), Babson et al. (1964), and Churchill (1965) studies, the samples were numerically small and highly selected. To obtain a larger, more representative sample, Kaelber and Pugh searched the records of four major Boston area hospitals to find 374 twin sets aged 6 to 16 years whose members had been tested on the same IQ measure within a 12-month period. Of these pairs, 44 were diagnosed monozygotic, 32 by blood typing or placental examination and 12 only by a physician's statement in the hospital birth record. They grouped the pairs into sets with small (< 300 gm) or large (-> 300 gm) birth weight differences. The only significant mean difference in their study, 5.76 IQ points favoring the heavier twin, was found among the 17 "MZ" pairs with large birth weight differences. But in a later erratum, Kaelber and Pugh (1969b) corrected the zygosity of two MZ pairs to DZ, which reduced this mean difference to a nonsignificant 5.13 points.

An investigation by Drillien (1970) involved 77 prospectively studied twin sets including 19 MZ pairs among which only 5 had birth weight differences exceeding 10%. There was no correspondence between IQ measured at ages 10 to 12 and birth weight in the 11 female pairs, but in the 8 male pairs, all the heavier members had higher IQs than their lighter mates. No significance tests were reported.

Like Kaelber and Pugh (1969a), Fujikura and Froehlich (1974) began their report by noting the shortcomings of earlier studies, such as use of retrospectively collected data, small samples, and questionable methods of zygosity diagnosis. Their paper was based on the prospectively studied NCPP twins, both Black and White, who had been tested on the Bayley Mental and Motor Examinations at 8 months and the Stanford-Binet at 4 years. To make the birth weight differences of lighter and heavier sets comparable, they computed the difference as a percentage of the heavier twins' birth weight: percent difference birth weight (PDB) = (A - B)/A x 100, where A is the weight of the heavier twin and B the weight of the lighter twin. Pairs were classed into small (<15%) or large (->15%) PDB groups with these groups further subdivided by race. The greatest differences were on the Stanford-Binet in the large-PDB groups, 1.86 and 4.73 IQ points, respectively, for the Blacks (np -- 15) and Whites (np = 11), but neither these nor any of the other differences were significant, including those for the Bayley tests.

Even though Fujikura and Froehlich (1974) criticized the small samples in earlier studies, O'Brien and Hay (1987) pointed out that Fujikura and Froehlich's own study had small samples of large PDB pairs and that their total sample also had a limited age range. It should be noted that Fujikura and Froehlich's was the third study of MZ intrapair variation done with the NCPP twins. These twins were tested at ages 4, 7, and 8. Analyzing the data from age 7, Melnick et al. (1978) found that in White twins, but not in Black, there was significantly greater intrapair variation in dichorionic than in monocho-rionic pairs. But Brown (cited in Welch et al., 1978), like Fujikura and Froeh-lich, found no differences in within-pair variability by chorion type in the data from age 4. Among all these analyses, by chorion type, large and small PDB groups, and within race, significant intrapair variability was shown in only one racial group at one age.

Wilson (1979) followed the growth of 10 MZ pairs from the Louisville Twin Study. The cotwin birth weight difference exceeded 750 grams in each of these pairs and averaged more than 1,000 grams. When tested at age 6, the mean intrapair IQ difference (favoring the heavier twin) was a nonsignificant 1.6 points.

Marsh's (1980) subjects were 46 reliably diagnosed MZ pairs whose ages were from 16 to 68 years and IQs from 104 to 135 when tested on Raven's Progressive Matrices. The mean IQ of the heavier cotwins was significantly higher (marginally) than that of the lighter cotwins by a one-tailed test, $t(45) = 1.68$, $p = .05$, intrapair birth weight differences accounting for 2.5% of the variance in IQ. Marsh stated that the magnitudes of the IQ and birth weight differences were unrelated but provided no correlations showing this.

O'Brien and Hay (1987) studied a sample of 100 MZ pairs (55 male and 45 female), 7 to 15 years of age to whom they administered six tests of verbal and performance skills. O'Brien and Hay performed analyses on the total sample, within gender, and within small (0 - 15%) and large (16% +) PDB groups. In the resulting large number of comparisons, the heavier cotwins scored significantly higher (5.8 points, $p < .0001$) than the lighter cotwins only in the male, large PDB group on one perceptual organization measure.

Munsinger (1977) initiated a controversy when he reanalyzed the data from six earlier studies (Babson et al., 1964; Churchill, 1965; Fujikura & Froehlich, 1974; Kaelber & Pugh, 1969a [ignoring the erratum, 1969b]; Scarr, 1969; Willerman & Churchill, 1967) in an attempt to show that MZ twin concordance estimates for IQ and the heritability estimates based on them are biased downward by sets in which the twin transfusion syndrome has seriously impaired the development of one cotwin resulting in both large birth weight and IQ differences. In subsequent critiques, however, Kamin (1978) and Marsh (1979) found that Munsinger's work was seriously in error. Kamin performed his own reanalysis on the raw data from five of the six studies Munsinger reanalyzed (excepting Kaelber & Pugh, 1969a). For a total of 97 MZ pairs, he found that the mean IQ of the heavier cotwins was about five points higher than that of the lighter cotwins. He also found that the magnitude of the intrapair IQ and birth weight differences were unrelated. Cotwins differing in birth weight by 500 grams differed in IQ no more than cotwins differing in birth weight by 300 grams or indeed by 50 grams. More-over, he showed that in the two studies with samples of same-sexed DZ pairs, the IQ advantage of the heavier cotwin was no greater among the MZ twins than among the DZ twins. After thus showing that the IQ advantage of the heavier cotwin is independent of both intrapair birth weight differences and zygosity, he asserted that the higher IQ of the heavier MZ cotwin could not be due to the twin transfusion syndrome.

James (1982) argued, however, that Kamin's (1978) assertion was unwarranted. He pointed out that fetal circulation is interconnected not only in pairs diagnosed with the transfusion syndrome but in about 90% of monocho-rionic pairs, and that the effects of shared circulation may not be limited to the lighter twin as Kamin assumed. In at least some cases where there is a large birth weight difference, the IQ of the heavier twin or of both twins may be detrimentally affected, and therefore IQ differences would not be expected to increase with birth weight differences. In other words, the heavier twin may have a lower IQ if its greater birth weight reflects a pathological condition such as edema rather than the better health normally associated with heavier birth weights. The issue, then, is not whether the IQ of the heavier or lighter twin is higher or lower but whether the IQ of the more seriously affected twin is higher or lower. This implies that determining which twin is more seriously affected must be done with criteria other than birth weight.

James (1982) also argued that Kamin's (1978) finding that IQ differences are unrelated to zygosity may have been a Type II error. Kamin's results were based on two studies (Churchill, 1965; Kaelber & Pugh, 1969a, 1969b) considered separately. Had he pooled samples, he may have found significantly greater cotwin differences among MZ pairs. James showed that combining Kamin's 97 MZ pairs with Kaelber and Pugh's 42 MZ pairs yielded a mean IQ advantage for the heavier twin of 4.44 points, whereas combining same-sexed DZ pairs from three studies (the two analyzed by Kamin plus Babson et al., 1964) yielded a mean IQ advantage of 1.76 points. Though the t value cannot be calculated because the raw data are not available, this may be a significant difference.

James's point is valid, but the pooling of results he suggested would be more properly done as a meta-analysis, the techniques for which were not well developed or well known at the time of James's article. A meta-analysis might show that a population difference too small to be significant in individual studies, such as those analyzed by Kamin, may be significant when studies are aggregated.

Marsh (1980), like Munsinger (1977) and Kamin (1978), also reanalyzed the data of earlier studies. For six intelligence measures used in five studies of MZ twins (Babson et al., 1964; Scarr, 1969; Willerman & Churchill, 1967; Shields's 1962 study of twins reared apart; and Marsh's own study), birth weight differences accounted for 3 to 21% of the variance in IQ, all significant proportions by one-tailed tests. Unfortunately, this reanalysis does little to clarify the issues of the relationship between intrapair IQ and birth weight differences, and whether intrapair IQ differences are related to zygos-ity. It may also have been inadvisable to include

Shields's study of twins reared separately because the other studies of IQ and birth weight differences, and most twin studies, are of twins reared together.

Table I0 summarizes the results of the 12 studies of intrapair intelligence and birth weight differences in MZ pairs. The results of studies analyzing only large PDB pairs are shown separately from studies in which both large and small PDB pairs were combined. Note that the Ns for all the large PDB groups are small. Note also that the samples overlap in three of the combined PDB studies (Churchill, 1965; Scarf, 1969; Willerman & Churchill, 1967). Of the six significant differences found in the combined PDB studies, five occur in the three studies with overlapping samples. This leaves only one significant difference among the remaining combined PDB samples.

Similarly, the samples are identical in two of the large PDB studies, Babson and Phillips (1973) being a follow-up to Babson et al. (1964). Of the five significant differences in the large PDB studies, four occur in this one sample. This leaves only one significant difference among the other large PDB samples. Thus, only two significant differences in all these studies occur in nonoverlapping samples. Most of the differences in the table are positive, however, favoring the heavier birth weight cotwin. This suggests that the heavier birth weight cotwin may have an IQ advantage, but it is small and requires a large sample to reliably detect it.

Several investigators have noted that cotwin differences are more often significant on performance than on verbal tests. Cotwin differences are significant on the performance measures in all four studies of combined PDB groups in Table 10, but three of these are the studies with overlapping samples. Performance tests were administered in only one large PDB study, O'Brien and Hay's (1987), where cotwin differences were significant only for the males on one measure. Like the other differences in the table, however, most of the differences on the performance tests are positive, favoring the heavier birth weight cotwin. This suggests that there may be a performance IQ advantage for the heavier birth weight cotwin, but one too small to show reliably without large samples.

Taken as a whole, these studies suggest that IQ differences between twins differing even greatly in birth weight are usually small. This implies that prenatal conditions are not an important source of MZ cotwin discordance biasing twin studies. As will be discussed in the conclusion below, Wilson (1979) presented evidence from large growth studies of twins that twins overcome early deficits resulting from intrauterine conditions.

It should be noted that if James's hypothesis that mutual circulation may affect either or both twins is correct, IQ differences are hidden in the large PDB groups because some of the differences favor the lighter rather than the heavier cotwins. This hypothesis should be tested. Doing so would require that the extent of fetal intercirculation be assessed for each MZ pair in a study; large initial samples would be necessary; and the study would have to be prospective from the twins' birth.

Physical Differences and Cognitive Differences Within MZ Pairs

Jensen (1987) has noted that the differences between MZ twins reared together are particularly useful because such intrapair differences result entirely from within-family environmental effects and are free of any genetic influence. By correlating MZ twin differences on cognitive tests with differences on other measures, variables influencing cognitive development can be identified. (The MZ differences also include measurement error, but this will be uncorrelated with other variables.) These variables may be either social or biological in nature and may be associated with either the pre- or postnatal environment. MZ differences may be used to identify variables that are important to cognitive development in the general population and also to investigate specific issues in twin analysis, including whether prenatal environmental factors are a source of bias in twin studies.

In the absence of direct measures of prenatal variables, the influence of prenatal factors can be inferred from correlations between MZ differences on physical and cognitive measures. If prenatal factors causing physical differences also cause cognitive differences, the differences should be positively correlated. For example, Newman, Freeman, and Holzinger (1937) found that among their 19 pairs of twins reared apart, differences in fingerprint ridges correlated .51 with differences in adult IQ (Jensen, 1987). This correlation suggests that the factors causing fingerprint ridges to differ also caused adult IQ to differ, and because dermatoglyphic patterns are set by the 20th week of gestation, these factors must have operated early in fetal development. Although this particular finding has not been confirmed in the Minnesota Stud.y of Twins Reared Apart (A. R. Jensen, personal communication; May 15, 1989), in general, if both twins have similar postnatal medical histories and adequate nutrition, it seems likely that the correlation of within-pair physical and cognitive differences may reflect the long-term effects of prenatal conditions on both physical and cognitive development. Note that these MZ within-pair correlations are WF correlations because they are computed from the differences between members of twin pairs. They are called within-pair or intrapair correlations here because these terms are used in the literature to refer to correlations based on MZ cotwin differences.

Within-Pair Analyses of Physical Differences and Cognitive Differences

For this review, I found four studies that correlated MZ twin within-pair physical differences and cognitive differences. Burks's (1940) analysis was based on 20 MZ pairs (14 male and 6 female) who had participated in the Harvard Growth Study between ages 6 and 18. Their intelligence had been tested and physical measurements taken at yearly intervals, some as few as 3 times, others as many as 11 times. They were administered standard IQ tests such as the Stanford-Binet and Otis, though the same tests were seldom given in successive years. This would not affect within-pair correlations, however, because both members of each pair took the same test in a given year. Correlations were computed between mean within-pair differences in IQ and mean within-pair differences on each physical measure. The averages were over the number of times each pair was observed and were therefore more reliable than differences based on single measurements.

The second column of Table 11 shows the correlations between the intra-pair differences in IQ and intrapair differences on five physical measures. Three of the intrapair correlations are substantially larger than the WI correlations generally found between physical and mental measures and larger than those found in Burks's (.1940) study, shown in Column 3 of the table, but all the intrapair correlations are nonsignificant. Note that although the correlations are nonsignificant, they may be biased upward because the averages they are based on include within-pair differences from early childhood, when the differences would be greater if children recover from prenatal deficits as they mature.

I performed within-pair analysis with the small number of MZ pairs for which Thurstone et al. (1955) reported complete test score and body size data, using the principal component scores for the six PMA tests described above as the intelligence measure. The within-pair correlation of the principal component scores with height was -.29 + -.42 (N = NP = 25) and with weight, .06 -+.57 (N = NP = 15). These correlations, which are neither significant nor consistent in direction, suggest that within-pair differences on the intelligence measures are unrelated to within-pair differences on the physical measures.

Jensen (1987) performed within-pair analysis on data for 82 MZ pairs from the Louisville Twin Study. Using the principal component scores noted earlier as his measure of intelligence, he obtained the multiple correlation between cotwin differences on the measure with differences in body size (height and weight) and also with differences in head size (length, width, and circumference). The multiple correlations for body size and head size, .04 and. 17, respectively, were both nonsignificant.

The fourth within-pair analysis (Johnson, 1990a) was done with height, weight, head circumference, and intelligence measures taken at ages 4 and 7 on 82 sets of identical twins from the NCPP. Correlations for the older age are shown in Table 12. (The discrepant Ns are due to missing data.) The correlations vary, with that for weight significant, that for head circumference negative. Note that this is the fourth study done with the NCPP data in which childhood test scores have been used to assess prenatal effects on intelligence. (The other three studies are Fujikua & Froehlich, 1974; Melnick et al., 1978; and O'Brien & Hay, 1987.) In this analysis and all the others, by chorion type and PDB groups, only inconclusive evidence of prenatal effects has been found.

Table 12 shows the results and mean N-weighted correlations based on them for the four within-pair analyses of intelligence and physical differences. Only one mean correlation is significant, but both mean correlations are positive, as are Jensen's multiple correlations. This suggests that a weak relationship between intrapair physical and intelligence differences may exist. A weak relationship implies that pre- or postnatal factors causing physical differences result in only small IQ differences and are not an important source of bias in twin studies. As discussed below, Wilson (1979) shows that twins largely recover from physical and cognitive deficits of prenatal origin.

Conclusions

Pleiotropy

IQ and head size. The review has shown that intelligence and head size are positively correlated in the populations sampled. The N-weighted mean partial correlation (controlling height) of IQ and head size is. 10. The estimate based on Van Valen's method and the results of the Willerman et al. (1989) study suggest that the correlation of intelligence and brain size is between. 30 and .40, which attributes about 10 to 15% of the variance in IQ to brain size. Although the results of the within-family analysis suggest that pleiotropy may play a role in the correlation, it appears to be a small one because only 26% of the WI correlation also occurs within families. This implies that the IQ/ head-size correlation is largely between families and arises primarily from environmental factors or common assortment of genes due to assortative mating. But this review has shown that the correlation occurs almost undiminished within SES levels, largely ruling out environmental factors related to social class. Moreover, the N-weighted mean partial correlation of. 10 indicates that the correlation persists when body size (height) is controlled, seemingly ruling out assortative mating. Given these results, it seems prudent to conclude that the primary source of the IQ/head-size correlation is a between-family factor that remains unidentified.

IQ and body size. The review has also shown that intelligence and body size are positively correlated in the populations sampled, with the N-weighted mean partial correlation (controlling age) of IQ with height about. 18 in children and .22 in adults. The mean age-controlled IQ-weight correlation is somewhat lower, about. 14 in children. (The age-controlled IQ-weight correlation was. 10 in one study of adults.) The within-family analyses suggest that pleiotropy may contribute to the IQ/body-size correlation. The pooled WF correlation is significant (marginally) and indicates that about 31% of the WI correlation is found within families. Clearly, if the within-family component of the correlation is due to pleiotropy, there must be a biochemical process initiated by at least one gene that affects both intelligence and body size. William J. Libby (personal communication; November 20, 1990) has noted that because there are many biochemical processes affecting growth and biological functioning, it is likely that at least some of them affect more than one trait. Thus, there may be many pleiotropisms, of which the correlation between intelligence and body size is just one.

The between-family component of the correlation must be due to environmental factors or common assortment of genes. Several of the studies reviewed provide evidence against environmental factors as the dominant source of the between-family component. First, Douglas et al. (1965) found that the WI correlation occurs undiminished within levels of sexual development. If the between-family component were largely due to socially patterned differences in health or nutrition affecting sexual development, the WI correlation should be much lower or zero within levels of sexual maturation. Second, Douglas et al., Broman et al. (1975), and Scott et al. (1956) found the WI correlation occurring about as strongly within SES levels as across SES levels. If the between-family component of the correlation were primarily due to factors varying among social classes, such as health, nutrition, or education, the WI correlation should be much lower or zero within SES levels.

Other lines of evidence suggest that the between-family component of the correlation involves genetic factors. Both IQ and body size are substantially heritable; thus, the correlation between them is likely to involve a genetic factor. Moreover, there is known positive assortative mating both for height and IQ. If there were positive cross-assortative mating for both traits, they would be correlated in the population through common assortment of genes. Support for the assortative mating hypothesis can be easily obtained by analyzing the correlation of height and IQ in spouses.

MZ Cotwin Concordance

In general, the evidence from the three kinds of studies reviewed suggests that prenatal conditions have only minor long-term effects on MZ cotwin concordance for IQ. The studies of chorionicity found at best. a weak relationship between chorion type and intrapair variability in IQ among MZ pairs studied over a wide age range, 4 years to adulthood. The studies of cotwin birth weight differences suggest that the heavier birth weight MZ cotwin has only a small IQ advantage over the lighter cotwin. The evidence that cotwin IQ differences and birth weight differences are correlated is at best weak. Such weak evidence would be obtained if, as James (1982) suggests, mutual circulation affects either twin. But this possibility has not yet been investigated. In the studies correlating cotwin intelligence differences and physical differences, they were also weakly related. This finding supports the previous one because birth weight and physical differences are likely to have the same prenatal origin.

Overall, these results suggest that prenatal conditions have little lasting effect on twin concordance. Wilson (1979) has further demonstrated this with physical traits. Studying the growth of 400 twins from birth through age 8, he found that although twins are 30% lighter and 17% shorter than singletons at birth, they recover from this initial size deficit. Lighter birth weight cotwins also attain the same height and weight, on average, as their heavier mates. Moreover, although MZ cotwins are less concordant than DZ cotwins in length and weight at birth, as noted above, they quickly become more concordant.

Wilson's finding, also noted earlier, that among 10 MZ pairs with cotwin birth weight differences exceeding 750 grams, IQ differed by only 1.6 points at age 8, led him to conclude that intelligence is highly buffered against prenatal nutritional deficit. Because of this buffering, each twin returns to its genetically programmed course of cognitive development despite handicapping intrauterine events. There is other evidence of buffering against deleterious prenatal effects on IQ. Naeye, Deiner, Dellinger, and Blanc (1969) compared the weight of eight organs of poor and nonpoor stillborn children and found the brain least affected by nutritional differences associated with poverty. Also, in a study of Dutch men examined at age 19 for military service, no difference in IQ was found between recruits whose mothers had or had not experienced famine during World War II while pregnant with them (Stein, Susser, Saenger, & Marolla, 1972; cited in Lerner & Libby, 1976, pp. 222-223).

Evidence indicates that, because of mutual circulation, there are much greater differences in the prenatal environments of MZ cotwins than of DZ cotwins. But the effects of these differing prenatal environments do not persist. The lighter birth weight MZ cotwin overcomes its initial disadvantage, and MZ pairs rapidly become more concordant than DZ pairs. The recovery process also appears to apply to the effects of health differences in the postnatal environment. Newman, Freeman, and Holzinger (1937) found health and physical differences uncorrelated with intelligence differences in their sample of 19 MZ pairs reared apart. Similarly, Loehlin and Nichols (1976) found differences in illness during infancy and childhood among MZ cotwins unrelated to test score differences in adolescence. In general, research suggests that factors affecting the physical development of twins, whether pre- or postnatal, have little long-term impact on MZ concordance for intelligence and are thus not likely to seriously bias estimates of genetic influence based on the twin method. This implies that intelligence differences among identical cotwins are largely a result of the postnatal social environment.

TABLE 1
Results From Studies of the Correlation Between Intelligence and Head Size

Study	Year	Sample	Intelligence measure
Adolescent and adult singletons			
Pearson	1906	Unselected	Ratings
Pearson	1906	University	Grades
Pearl	1906	Soldiers	Ratings
Murdoch & Sullivan	1923	Unselected	IQ tests
Reid & Mulligan	1923	University	Grades
Sommerville	1924	University	IQtest
Passingham	1979	Unselected	IQtest
Suzanne	1979	Unselected	Tests
Adolescent and adult twins			
Jensen r	1987	Unselected	IQtests
Singleton children			
Robinow	1968	?	IQ test
Weinberg et al.	1974	Unselected	IQ test
Broman et al.	1975	Unselected	IQ test
Johnson	1990b	Unselected	Tests
Johnson	1990b	Unselected	Tests

Study	N	r IC	r IC.H	95% CI	Controlled variables
Adolescent and adult singletons					
Pearson	4,486	.11	.05	-+.03	AS
Pearson	1,011	.11	.12	-+.06	AS
Pearl	935	.14	.09	-+.06	AS
Murdoch & Sullivan	595	.22	.18	-+.08	AS
Reid & Mulligan	449	.08	.07	-+.09	A S H
Sommerville	105	.08	.10	-+.19	A S
Passingham	415	.13	.08	-+.10	A H

Suzanne	2,071	.22	.18	-+.04	A S H
		Adolescent and adult twins			
Jensen	286	.30b			A S
r		.14c	.10	-+.02	
		Singleton children			
Robinow	300	.18	.09	-+.11	A S
Weinberg et al.	334	.35	.29	-+.11	S
Broman et al.	21,759	.14d	.12	-+.01	A
Johnson	3,236	.15e	.13	-+.03	A S
Johnson	3,202	.16e	.13	-+.03	A S

Note. A = age; C = external cranial size; CI = confidence interval; H = height; I = intelligence; S = sex. aSee text for explanation of how $r_{IC.H}$ was computed for each study. bMultiple correlation of unrotated first principal component scores with three head measures, height, width, and circumference (p < .01). cExcludes Jensen's multiple R. an-weighted mean of two groups. eMean for two tests.

TABLE 2
Correlations Reported by Broman et al. (1975) of Stanford-Binet IQ With Head Circumference at Age 4 in the Total White and Black Groups and Within SES Levels

SES	White	Black
Lowest 25%	.13 -+.06	.09 -+.03
Middle 50%	.09 -+.03	.11 -+.02
Highest 25%	.12 -+.03	.13 -+.06
Total group	.15 -+.02	.12 -+.02

Note. Confidence intervals are 95%.

TABLE 3
Results From Within-Family Analyses of the Intelligence/Head-Size Correlation

Subject	r WF	97.5% CI	r WI	97.5% CI
		Same-sexed DZ twins		
		Johnson (1990a)		
	(N =NP = 56)		(N = 2NP = 112)	
4 years	.08	+.27	.26	+.15
7 years	.12	+.27	.26	+.15
		Jensen (1987)		
	(N = NP = 61)		(N = 2NP = 286a)	
Multiple R	.28b		.30b	
Full siblings				
		Johnson (1990b)		
5-year cohort				
	(N= NP = 153)		(N= 2NP = 306)	

CPM	.01	+.16	.22	+.11
PPVT	.16	+.16	.15	+.11
10-year cohort				
	(N= NP = 338)		(N = 2NP = 676)	
SPM	.00	+.11	.12	+.08
PPVT	.08	+.11	.22	+.08
	.06c	+.08	.19c	+.06

Note. Confidence intervals are one-tailed. CI = confidence interval; DZ = dizygotic; MZ = monozygotic; WI = within individual; WF = within family. a For the total sample of MZ and same-sexed DZ twins in the study. b Multiple correlation of unrotated first principal component scores with three head measures (height, width, and circumference) for the total sample of MZ and same-sexed DZ twins in the study; the WI multiple R is significant (p < .01). cExcludes Jensen's multiple R.

TABLE 4
Results From Studies of the Correlation Between Intelligence and Height in Children

Study	Year	Sample	Intelligence measure
Singletons			
Murdoch & Sullivan	1923	Unselected	IQ tests
Baldwin	1925	Gifted	IQ test
Dearborn et al.	1938	Unselected	IQ tests
Scottish Council	1953	Unselected	IQ test
Bayley	1956	Unselected	IQ tests
Scott	1962	Unselected	IQ tests
Douglas et al.	1965	Unselected	Test
Wilson et al.	1986	Unselected	IQ test
Nagoshi & Johnson	1987	Unselected	Tests
Johnson	1990b	Unselected	Tests
Twins			
Burks	1940	Unselected	IQ tests
Thurstone et al.	1955	Unselected	IQ tests
Jensen	1987	Unselected	IQ tests
Young children			
Weinberg et al.	1974	Unselected	IQ test
Broman et al.	1975	Unselected	IQ test
Johnson	1990b	Unselected	Tests

Study	N	r_{IH}	$r_{IH.A}$	95% CI
Murdoch & Sullivan	597		.14	-+.08
Baldwin	594		.21a	-+.08
Dearborn et al.	1,048		.20a	-+.06
Scottish Council	6,490		.25a	-+.02
Bayley	40		.37a	-+.32
Scott	4,259		.23a	-+.03

Douglas et al.	2,864		.12a	-+.04
Wilson et al.	13,887	.14a		-+.02
Nagoshi & Johnson	1,326		.08a	-+.05
Johnson	3,279		.08b	-+.05
		Twins		
Burks	21	.17		-+.46
Thurstone et al.	118		.23	-+.18
Jensen	286		.08c	
r		.14d	.18d	-+.02
		Young children		
Weinberg et al.	334	.21		-+.11
Broman et al.	21,854		.07a	-+.01
Johnson	3,236		.09a	-+.03

Note. A = age; CI = confidence interval; H = height; I = intelligence. an-weighted mean of two groups. bMean of correlations at ages 5 and 10 and during adolescence. CNonsignificant multiple correlation of IQ with height and weight. $_d$ $r_{IH.A}$ excludes Dearborn et al. because their sample included heterogeneous groups, possibly inflating the correlation; Bayley, whose correlation was unequal to the others by a test of homogeneity (Hedges & Olkin, 1985); and Jensen, who computed a multiple R.

TABLE 5

Results From Studies of the Correlation Between Intelligence and Height in Adults

Study	Year	Sample	Intelligence measure
		Singletons	
Lee et al.	1902	University	Grades
Reid & Mulligan	1923	University	Grades
Sommerville	1924	University	IQ test
Scott et al.	1956	Unselected	IQ tests
Husen	1959	Unselected	Test
Husen	1959	Unselected	Test
Udjus	1964	Unselected	Test
Passingham	1979	Unselected	IQ test
Suzanne	1979	Unselected	Tests
Nagoshi & Johnson	1987	Unselected	Tests
Teasdale et al.	1990	Unselected	Tests
		Twins	
Husen	1959	Unselected	Test
r			

Study	N	r_{IH}	$r_{IH.A}$	95% CI
Lee et al.	1,011	-.01		-+.06
Reid & Mulligan	449		.05	-+.09
Sommerville	105		.16	-+.19
Scott et al.	278	.21a		-+.12
Husen	2,250		.22	-+.04
Husen	5,000		.20	-+.03
Udjus	4,458		.16	-+.03
Passingham	415		.13a	-+.10

Suzanne	2,071	.18		-+.04
Nagoshi & Johnson	2,725		.14a	-+.04
Teasdale et al.	71,528	.22		-+.00
Husen	682		.24a	-+.08
r		.19b	.22c	-+.01d

TABLE 6
Results From Studies of the Correlation Between Intelligence and Weight

Study	Year	Sample	Intelligence measure
Singleton children			
Murdoch & Sullivan	1923	Unselected	IQ tests
Baldwin	1925	Gifted	IQ Test
Scottish Council	1953	Unselected	IQ test
Scott	1962	Unselected	IQ tests
Johnson	1990b	Unselected	Tests
Twin children			
Thurstone et al.	1955	Unselected	IQ tests
Burks	1940	Unselected	IQ tests
r			
Young children			
Weinberg et al.	1974	Unselected	IQ test
Broman et al.	1975	Unselected	IQ tests
Johnson	1990b	Unselected	Tests
Adults			
Sommerville	1924	University	IQ test

Study	N	r_{IW}	$r_{IW.A}$	95% CI
Singleton children				
Murdoch & Sullivan	595		.16	-+.08
Baldwin	594		.04a	-+.08
Scottish Council	6,490		.19a	-+.02
Scott	4,259		.13a	-+.03
Johnson	3,277		.07b	-+.05
Twin children				
Thurstone et al.	82		.24	-+.22
Burks	21	-.02		-+.46
r			.14c	-+.02
Young children				
Weinberg et al.	.334		.11	-+.11
Broman et al	21,853		.15a	-+.01
Johnson	3,346		.06a	-+.03

	Adults		
Sommerville	105	.10	-+.19

Note. A = age; CI = confidence interval; I = intelligence; W = weight.

[a] n-weighted mean of two groups. bMean of correlations at ages 5 and 10 and during adolescence. cExcludes Baldwin by a test of homogeneity.

TABLE 7

Correlations Reported by Broman et al. (1975) of Stanford-Binet IQ With Height and Weight at Age 4 in the Total White and Black Groups and Within SES Levels

	Whites		Blacks	
SES	Height	Weight	Height	Weight
Lowest 25%.	12 -+.06	.11 -+.06	.14 -+.03	.15 -+.03
Middle 50%.	06 -+.03	.07 -+.03	.10 -+.02	.13 -+.02
Highest 25%	.08 -+.03	.10 -+.03	.12 -+.06	.10 -+.06
Total group	.10 -+.02	.10 -+.02	.13 -+.02	.15 -+.02

Note. Confidence intervals are 95%.

TABLE 8

Results From Within-Family Studies of the Intelligence/Body-Size Correlation

Study	Year	NP	rwF	97.5% CI	rwi	97.5% CI
			Full siblings			
Laycock & Caylor	1964	81	ns[a]			
Jensen	1980	2,396	.02	+.04		
Nagoshi & Johnson	1987	467 AEA	.06	+.09	.10**	+.07
		144 AJA	-.02	+.17	.05	+.11
Johnson	1990b	310	.07*b	+.11	.10***b	+.08
		153	.06	+.16	.15***	+.11
			DZ twins			
Thurstone et al.	1955	34	.27	+.35	.23c	+.18
Jensen	1987	61	.20d	.08a		
Johnson	1990a	56	.27*e	+.27	.20e	+.15
r			.04f	+.02	.12f	+.05

Note. Confidence intervals are one-tailed. AEA = Americans of European ancestry; AJA = Americans of Japanese ancestry; CI = confidence interval; DZ = dizygotic; NP = number of pairs; ns = nonsignificant; WF = within family; WI = within individual. aNonsignificant difference between sibs. bMean of correlations at ages 5 and 10 and during adolescence. cFor the total number of MZ and same-sexed DZ twins in the study. dNonsignificant multiple correlation of first principal component scores with height and weight for the total number of MZ and same-sexed DZ twins in the study. eMean of correlations at ages 4 and 7. fExcludes Jensen's multiple correlation.

p < .025, one-tailed. **p < .005, one-tailed. ***p < .0005, one tailed.

TABLE 9

Results From Studies Comparing the Intrapair Variability in

Intelligence of Monochorionic and Dichorionic MZ Twins

Study	Year	N_{Mc}	N_{Dc}
Breland	1974	117	365
Brownb	1977	55	44
Melnick et al.	1978	23	9
		30	24
Welch et al.	1978	20	12
Rose et al.	1981	17	15

Study	Age	Result[a]	Test
Breland	16 yrs	+ (ns)	NMSQT
Brownb	4yrs	(ns)c	Stanford-Binet
Melnick et al.	7 yrs, White	+ (**)	WISC
	7 yrs, Black	+ (ns)	WISC
Welch et al.	18 mos	+ (ns)	Bayley
Rose et al.	Adults	(ns)d	WAIS Vocab
		+ (*)e	WAIS Blocks

Note. DC = clichorionic; MC = monochorionic; MZ = monozygotic; NMSQT = National Merit Scholarship Qualifying Test; ns = nonsignificant; WAIS = Wechsler Adult Intelligence Scale; WISC = Wechsler Intelligence Scale for Children.

[a] Plus signs indicate that intrapair variability was greater among monochorionic than dichorionic twins. bSample overlapped with Melnick et al's. cDirection of the difference was not stated in the source article. dIntrapair variability was equal for the monochorionic and dichorionic twins. eDifference significant, but level unspecified.

p < .05. **p < .01.

TABLE 10
Mean IQ Differences Between Heavier and Lighter
Birth Weight MZ Cotwins

Study	Year	NP
	Combined PDB pairs	
Churchill	1964	22
Willerman & Churchill	1967	27
Scarr	1969	52
Drillien	1970	11 females
8 males		
Marsh	1980	46
	Large PDB pairs	
Babson et al.	1964	9
		9
Babson &	1973	9
Phillips		9
Kaelber & Pugh	1969	15
Fujikura	1974	15 White

& Froehlich		11 Black	
Wilson	1979	10	
O'Brien & Hay	1987	14 males	
O'Brien & Hay	1987	15 females	

	Intelligence Measure			
	Verbal	Performance	Full scale IQ	Test
Combined PDB pairs				
Churchill	+1.9	+6.2***	+4.3***	WISC
Willerman & Churchill	+4.4*	+5.3***		WISC
Scarr		+6.9***		DAP
Drillien			a	?
			+b	?
		+*c		Raven
Large PDB pairs				
Babson et al.			+6.6**	S-Bb
		+2.7*		PPVT
Babson & Phillips		+8.7**		WISC-V
		+7.4**		PPVT
Kaelber & Pugh			+5.1	Group
Fujikura & Froehlich			+4.7	S-B
			+1.9	S-B
Wilson		+1.6		WISC (?)
O'Brien & Hay		-2.9		PPVT
		+5.8***		Blocks
		0.0		Cubes
		-1.0		PPVT
O'Brien & Hay		+1.1		Blocks
		+0.9		Cubes

Note. Blocks = Block Design; Cubes -- Knox Cubes; DAP = Draw-a-Person; MZ = mono-zygotic; NP = number of pairs; PDB = percent difference in birth weight; PPVT = Peabody Picture Vocabulary Test; S-B = Stanford-Binet; WISC(-V) = Wechlser Intelligence Scale for Children (Verbal).

[a] Small IQ difference, but no information on the direction or significance of the difference was reported. [b] IQ difference favored the heavier cotwin, but no information on significance was given. cThe mean IQ difference was not reported, but it favored the heavier cotwin.

*p < .05. **p < .01. ***p < .001.

TABLE 11

Results From Burks's (1940) Within-Pair Analysis of Intelligence Differences and Physical Differences in MZ Cotwins

Physical measures	'Mean within-pair difference correlation (N = NP = 20)	WI correlation (N = NP = 21 a)
Height	.47 + .49b	.17 + .46c

Trunk length	.40 + .49	.08 + .46
Leg length	.11 + .49	.29 + .46
Iliac	.41 + .49	-.11 + .46
Weight	.12 + .49	-.02 + .46

Note. NP = number of pairs; WI = within individual. aNs are as presented in Burks's paper and are unequal presumably because of missing data. b97.5% one-tailed confidence interval. c95% two-tailed confidence interval.

TABLE 12
Summary of Results From Studies of MZ Within-Pair Cognitive Differences and Physical Differences

Study	Year	Height	Weight	Head circ	Ns
		Correlations			
Burks	1940.	.47a	.12a		20 20
Thurstone et al.	1955	-.29	.06		25 15
Johnson	1990a	.10	.24*	-.03	80 79 82
r		.09	+.17b	.20 +.18b	
		Multiple correlations			
Jensen	1987	.04c		.17d	82 82

Note. NP = number of pairs. a Correlations are for within-pair differences averaged over from 3 to 11 testings at yearly intervals. bConfidence intervals are 97.5% one-tailed. cMultiple correlation of first principal component scores with height and weight. dMultiple correlation of first principal component scores with three head measures (height, width, and circumference).

*p < .05, one-tailed.

REFERENCES

Allen, G., & Kallmann, F. J. (1962). Etiology of mental subnormality in twins. In F. J. Kallmann (Ed.), Expanding goals of genetics in psychiatry (pp. 174-191). New York: Grune.

Babson, S. G., Kangas, J., Young, N., & Bramhall, J. L. (1964). Growth and development of twins of dissimilar size at birth. Pediatrics, 30, 327-333.

Babson, S. G., & Phillips, D. $. (1973). Growth and development of twins dissimilar in size at birth. The New England Journal of Medicine, 289,937-940.

Bailey, J. M., & Horn, J. M. (1986). A source of variance in IQ unique to the lower-scoring monozygotic (MZ) cotwin. Behavior Genetics, 16, 509-516.

Baldwin, B. T. (1925). Anthropometric measurements. In L. M. Terman (Ed.), Genetic studies of genius: Vol I. Mental and physical traits of a thousand gifted children (pp. 135-171). Stanford, CA: Stanford University Press.

Bayley, N. (1956). Individual patterns of development. Child Development, 27, 45-74.

Breland, N. S. (1974). A test of primary bias in twin studies with respect to measured ability. Behavior Genetics, 4, 101-109.

Broman, S., Nichols, P. L., & Kennedy, W. A. (1975). Preschool IQ: Prenatal and early developmental correlates. Hillsdale, N J: Erlbaum.

Brown, B. (1977). Placentation effects on birth weight and IQ in MZ twins. Paper presented at the meeting of the Society for Research in Child Development, New Orleans, LA.

Bulmer, M. G. (!970). The biology of twinning in man. Oxford, England: Clarendon.

Burks, B. S. (1940). Mental and physical development patterns of identical twins in relation to organismic growth theory. In G. M. Whipple (Ed.), Intelligence: Its nature and nurture. Thirty-ninth yearbook of the National Society for the Study of Education, Part H, (pp. 85-96). Bloomington, IL: Public School Publishing Co.

Churchill, J. A. (1965). The relationship between intelligence and birth weight in twins. Neurology, 15, 341-347.

Dearborn, W. F., Rothney, J. W. M., & Shuttleworth, F. K. (1938). Data on the growth of public school children. Monographs of the Society for Research on Child Development, 31, 1.

Douglas, J. W. B., Ross, J. M., & Simpson, H. R. (1965). The relation between height and measured educational ability in school children of the same social class, family size, and stage of sexual development. Human Biology, 37, 178-186.

Drillien, C. M. (1970). The small for date infant: Etiology and prognosis. Pediatric Clinics of North America, 17, 9-24.

Fujikura, T., & Froehlich, L. A. (1974). Mental and motor development in mono-zygotic cotwins with dissimilar birth weights. Pediatrics, 53, 884-889.

Hedges, L. V., & Olkin, I. (1985). Statistical methods for meta-analysis. Orlando, FL: Academic Press.

Husen, T. (1959). Psychological twin research I. A methodological study. Stockholm: Almqvist & Wiksell.

James, W. H. (1982). The IQ advantage of the heavier twin. British Journal of Psychology, 73, 513-517.
Jensen, A. R. (1980a). Bias in mental testing. New York: The Free Press.

Jensen, A. R. (1980b). Uses of sibling data in educational and psychological research. American Educational Research Journal, 17, 153-170.

Jensen, A. R. (1987, June). Twins: The puzzle of non-genetic variance. Paper presented at the 17th annual meeting of the Behavior Genetics Association, Minneapolis.

Jensen, A. R. (1990). Physical correlates of human intelligence. Unpublished manuscript.

Jerison, H. J. (1982). The evolution of biological intelligence. In R. J. Sternberg (Ed.), Handbook of human intelligence (pp. 723-791). Cambridge, England: Cambridge University Press.

Johnson, F. W. (1990a). Within-family analysis of cognitive and physical variables in twins. Manuscript submitted for publication.

Johnson, F. W. (1990b). Within-family analysis of cognitive and physical variables in siblings and twins. Unpublished doctoral dissertation, University of California, Berkeley.

Kaelber, C. T., & Pugh, T. F. (1969a). Influence of intrauterine relations on the intelligence of twins. The New England Journal of Medicine, 280, 1030-1034.

Kaelber, C. T., & Pugh, T. F. (1969b). Erratum. The New England Journal of Medicine, 281,332.

Kamin, L. (1978). Transfusion syndrome and the heritability of IQ. Annals of Human Genetics, 42, 16 I-I 71.

Laycock, F., & Caylor, J. S. (1964). Physiques of gifted children and their less gifted siblings. child Development, 35, 63-74.

Lee, A., Lewenz, M. A., & Pearson, K. (1902). On the correlation of the mental and physical characters in man. Part II. Proceedings of the Royal Society, 71,106114.

Lerner, I. M., & Libby, W. J. (1976). Heredity, evolution, and society (2nd ed.). San Francisco: W. H. Freeman.

Loehlin, J. C., & Nichols, R. C. (1976). Heredity, environment, and personality. Austin: University of Texas Press.

Marascuilo, L. A., & Levin, J. R. (1983). Multivariate statistics in the behavioral sciences: A researcher's guide. Monterey, CA: Brooks/Cole.

Marsh, R. W. (1979). The identical-twin transfusion syndrome and intelligence: A second look. British Journal of Psychology, 70, 413-417.

Marsh, R. W. (1980). The significance for intelligence of differences in birth weight and health with monochozygotic twin pairs. British Journal of Psychology, 71, 63-67.

Matheny, A. P., Wilson, R. S., & Dolan, A. B. (1976). Relations between twins' similarity of appearance and behavioral similarity: Testing an assumption. Behavior Genetics, 6, 343-352.

Melnick, M., Myrianthopoulos, N. C., & Christian, J. C. (1978). The effects of chorion type on variation in IQ in the NCPP twin population. American Journal of Human Genetics, 30, 425-433.

Munsinger, H. (1977). The identical twin transfusion syndrome: A source of error in estimating IQ resemblance in heritability. Annals of Human Genetics, 40, 307321.

Murdoch, K., & Sullivan, L. R. (1923). A contribution to the study of mental and physical measurement in normal children. American Physical Education Review, 28, 209-215,276-280, 328-330.

Naeye, R. L., Deiner, M. M., Dellinger, W. S., & Blanc, W. A. (1969). Urban poverty: Effects of prenatal nutrition. Science, 166, 1026.

Nagoshi, C. T., & Johnson, R. T. (1987). Between- vs. within-family analysis of the correlation of height and intelligence. Social Biology, 34, 110-113.

Newman, J., Freeman, F., & Holzinger, K. (1937). Twins: A study of heredity and environment. Chicago: University of Chicago Press.

O'Brien, P. J., & Hay, D. A. (1987). Birthweight differences, the transfusion syndrome and the cognitive development of monozygotic twins. Acta Geneticae Medicae et Gemellogiae, 36, 181-196.

Passingham, R. E. (1979). Brain size and intelligence in man. Brain, Behavior, and Evolution, 16, 253-270.

Paterson, D. G. (1930). Physique and intellect. New York: Century.

Pearl, R. (1906). On the correlation between intelligence and the size of the head. Journal of Comparative and Neurological Psychology, 16, 189-199.

Pearson, K. (1906). On the relationship of intelligence to size and shape of the head, and to other physical and mental characters. Biometrika, 5, 105-146.

Plomin, R., Willerman, L., & Loehlin, J. C. (1976). Resemblance in appearance and the equal environments assumption in twin studies of personality traits. Behavior Genetics, 6, 43-52.

Price, B. (1950). Primary biases in twin studies: A review of prenatal and natal difference-producing factors in monozygotic pairs. American Journal of Human Genetics, 2, 293-352.

Rausen, S. R., Masako, S., & Strong, L. (1965). Twin transfusion syndrome. Journal of Pediatrics, 66, 613-628.

Reid, R. W., & Mulligan, J. H. (1923). Relation of cranial capacity to intelligence. Journal of the Royal Anthropological Institute, 53, 322-331.

Robinow, M. (1968). The relationship between stature, head circumference, and brain size during childhood and adolescence. Paper presented at the Child Development Section of the Academy of Pediatrics.

Rose, R. J., Uchida, I. A., & Christian, J. C. (1981). Placentation effects on cognitive resemblance of adult monozygotes. In L. Gedda, P. Parisi, & W. E. Nance (Eds.). Twin research 3: Intelligence, personality, and development (pp. 35-41). New York: Alan R. Liss.

Scarr, S. (1969). Effects of birth weight on later intelligence. Social Biology, 16, 249-256.

Scarr, S., & Carter-Saltzman, L. (1979). Twin method: Defense of a critical assumption. Behavior Genetics, 9, 527-542.

Scott, E. M., Illsley, B. A., & Thomson, A.M. (1956). A psychological investigation of primigravidae II. Maternal social class, age, physique and intelligence. Journal of Obstetrics and Gynaecology of the British Empire, 63, 338-343.

Scott, J. A. (1962). Intelligence, physique, and family size. British Journal of Social Preventative Medicine, 16, 165-173. London: University Press.

Scottish Council for Research in Education. (1953). Social implications of the 1947 Scottish Mental Survey. London: Oxford University Press.

Shields, J. (1962) Monozygotic twins. London: Oxford University Press.

Sommerville, R. C. (1924). Physical, motor, and sensory traits. Archives of Psychology, 75, 1-108.

Strong, S. J., & Corney, G. (1967). The placenta in twin pregnancy. London: Pergamon.

Suzanne, C. (1979). On the relation between psychometric and anthropometric traits. American Journal of Physical Anthropology, 51,421-423.

Tanner, J. M. (1966). Galtonian eugenics and the study of growth: The relation of body size, intelligence test score, and social circumstances in children and adults. The Eugenics Review, 58, 122-135.

Tanner, J. M. (1969). Relation of body size, intelligence test scores, and social circumstances. In P. Mussen, J. Langer, & M. Covington (Eds.), Trends and issues in child development psychology (pp. 182-201). New York: Holt, Rinehart, & Winston.

Teasdale, T. W., Owen, D. R., & Sorensen, T. I. A. (1990). Intelligence and educational level in extremely short and extremely tall adult males. Manuscript submitted for publication.

Thurstone, T. G., Thurstone, L. L., & Strandskov, H. H. (1955). [A psychological study of twins 2. Scores of one hundred and twenty-five pairs of twins on fifty-nine tests]. Unpublished raw data.

Udjus, L. G. (1964). Anthropological changes in Norwegian men in the twentieth century. Oslo: Universitatsforlaget.

Van Valen, L. (1974). Brain size and intelligence in man. American Journal of Physical Anthropology, 40, 417-424.

Weinberg, W. A., Dietz, S. G., Penick, E. C., & McAlister, W. H. (1974). Intelligence, reading achievement, physical size, and social class. Journal of Pediatrics, 85, 482-489.

Welch, P., Black, K. N., & Christian, J. C. (1978). Placental type and Bayley mental development scores in 18-month-old twins. In W. E. Nance (Ed.), Twin research: Psychology and methodology (pp. 145-149). New York: Alan R. Liss.

Whipple, G. M. (1914). Manual of mental and physical tests Part I: Simpler processes. Baltimore: Warwick & York.

Willerman, L., & Churchill, J. A. (1967). Intelligence and birth weight in identical twins. Child Development, 38, 623-629.

Willerman, L., Schultz, R., Rutledge, H. N., & Bigler, E. (in press). Magnetic resonance imaged brain structure and intelligence. Intelligence.

Wilson, D. M., Hammer, L. D., Duncan, P.M., Dombush, S. M., Ritter, P. L., Hintz, R. L., Gross, R. T., & Rosenfeld, R. G. (1986). Growth and intellectual development. Pediatrics, 78, 646-650.

Wilson, R. S. (1979). Twin growth: Initial deficit, recovery, and trends in concordance from birth to nine years. Annals of Human Biology, 6, 205-220.

Received January 8, 1991

Preparation of this article was supported by the Institute for the Study of Educational Differences.

Requests for reprints should be sent to Fred W. Johnson, Educational Psychology Office, 4511 Tolman Hall, University of California, Berkeley, CA 94720.

~~~~~~~~

By FRED W. JOHNSON, School of Education,University of California, Berkeley
~~~~~~~~