IQ: THE RANK ORDERING OF THE WORLD / 153

affiliation, have heritabilities of zero. The claim of biological determinists has been that the heritability of IQ is about 80 percent. How do they arrive at this figure?

Estimating the Heritability of IQ

All genetic studies are studies of the resemblances of relatives. If a trait is heritable, that is, if different genotypes have different average performances, then relatives ought to resemble each other more closely than unrelated persons do, since relatives share genes from common ancestors. Brothers and sisters ought to be more like each other than aunts and nephews, who ought to be more similar than totally unrelated people. The standard measure of similarity between things that vary quantitatively is their correlation, which measures the degree to which larger values for one variable go together with larger values of a second variable, and smaller values with smaller values. The correlation coefficient, r, ranges from +1.0 for perfect positive correlation, through zero for no relationship, to −1.0 for perfect negative correlation. So, for example, there is a positive correlation between father's income and child's years of schooling. Richer fathers have better-educated children while poorer fathers have less-educated children, on the average. The correlation is not perfect, since some families produce children who go to graduate school, but it is positive. In contrast, in the United States there is a negative correlation between family income and the number of visits per year to hospital emergency rooms. The lower your income, the more likely you are to use the emergency room as a medical service instead of a private doctor.

One important point about correlation is that it measures how two things vary together but does not measure how similar their average levels are. So the correlation between the heights of mothers and their sons could be perfect in that taller mothers had the taller sons and shorter mothers had the shorter sons, yet all the sons could be taller than all the mothers. Covariation is not the same as identity. The significance of this fact for the heritability of IQ and its meaning is considerable. Suppose a group of fathers had IQs of 96, 97, 98, 99, 100, 101, 102, and 103, while their daughters, separated from their fathers at birth and raised by foster parents, had IQs respectively of 106, 107, 108, 109, 110, 111, 112, and 113. There is a perfect correspondence between the IQs of fathers and daughters, and we might judge the character to be perfectly heritable because, knowing a father's IQ, we could tell without error which of the daughters was his. The correlation is, in fact, r = +1.0, yet the daughters are ten points above their fathers in IQ, so the experience of being raised by foster parents had a powerful effect. There is thus no contradiction between the assertion that a trait is perfectly heritable and the assertion that it can be changed radically by environment. As we shall see, this is not a hypothetical example.

Second, a correlation between two variables is not a reliable guide to causation. If A and B are correlated, one may be the cause of the other, they may both
be the consequence of a common cause, or they may be entirely accidentally related. The number of cigarettes smoked per day is correlated with the chance of lung cancer because smoking is a cause of lung cancer. The floor area of a person's house and the average age to which he or she lives are positively correlated not because living in a big house is conducive to health but because both characteristics are a consequence of the same cause—high income. For that matter, the distance of the Earth from Halley's comet and the price of fuel are negatively correlated in recent years because one has been decreasing while the other increased, but for totally independent reasons.

In general, heritability is estimated from the correlation of a trait between relatives. Unfortunately, in human populations two important sources of correlation are conflated: Relatives resemble each other not only because they share genes but also because they share environments. This is a problem that can be circumvented in experimental organisms, where genetically related individuals can be raised in controlled environments, but human families are not rat cages. Parents and their offspring may be more similar than unrelated persons because they share genes but also because they share family environment, social class, education, language, etc. To solve this problem, human geneticists and psychologists have taken advantage of special circumstances that are meant to break the tie between genetic and environmental similarity in families.

The first circumstance is adoption. Are particular traits in adopted children correlated with their biological families even when they have been separated from them? Are identical (i.e., monzygotic, or one-egg) twins, separated at birth, similar to each other in some trait? If so, genetic influence is implicated. The second circumstance holds environment constant but changes genetic relationship. Are identical twins more alike than fraternal (i.e., dizygotic, or two-egg) twins? Are two biological brothers or sisters (sibs) in a family more alike than two adopted children in a family? If so, genes are again implicated because, in theory, identical twins and fraternal twins have equal environmental similarity but they are not equally related genetically.

The difficulty with both these kinds of observations is that they only work if the underlying assumptions about environment are true. For the adoption studies to work, it must be true that there is no correlation between the adopting families and the biological families. There must not be selective placement of adoptees. In the case of one-egg and two-egg twins, it must be true that identical twins do not experience a more similar environment than fraternal twins. As we shall see, these problems have been largely ignored in the rush to demonstrate the heritability of IQ.

The theory of estimating heritability is very well worked out. It is well known how large samples should be to get reliable estimates. The designs of the observations to avoid selective adoptions, to get objective measures of test performance without bias on the part of the investigator, to avoid statistical artifacts that may arise from unrepresentative samples of adopting families, are all well laid out in textbooks of statistics and quantitative genetics. Indeed, these theories are constantly put into practice by animal breeders who would be unable to have their research reports published in genetics journals unless they adhered strictly to the standard methodological requirements. The record of psychometric observations on the heritability of IQ is in remarkable contrast. Inadequate sample sizes, biased subjective judgments, selective adoption, failure to separate so-called "separated twins," unrepresentative samples of adoptees, and gratuitous and untended assumptions about similarity of environments are all standard characteristics in the literature of IQ genetics. There has ever been, as we shall see, massive and influential fraud. We will review in some detail the state of psychometric genetic observations—not simply because it calls into question the actual heritability of IQ, but because it raises the far more important issue of why the canons of scientific demonstration and credibility should be so radically different in human genetics than in the genetics of pigs. Nothing demonstrates more clearly how scientific methodology and conclusions are shaped to fit ideological ends than the sorry state of the heritability of IQ.

The Cyril Burt Scandal

The clearest evidence, by far, for the genetic determination of IQ was the massive life's work of the late Sir Cyril Burt. In 1969 Arthur Jensen quite correctly referred to Burt's work as "the most satisfactory attempt" to estimate the heritability of IQ. When Burt died, Jensen referred to him as "a born nobleman," whose "larger, more representative samples than any other investigator in the field has ever assembled" would secure his "place in the history of science." Hans Eysenck wrote that he drew "rather heavily" on Burt's work, citing "the outstanding quality of the design and the statistical treatment in his studies." 14

The Burt data seemed so impressive for a number of very good reasons. First, one of the simplest ways, at least in theory, of demonstrating the heritable basis of a trait is to study separated identical twins. The separated twin pairs have identical genes, and they are assumed not to have shared any common environment. Thus, if they resemble one another markedly in some respect, the resemblance must be due to the only thing they share in common: their identical genes. The largest IQ study of separated identical twins ever reported, supposedly based on fifty-three twin pairs, was that of Cyril Burt. The IQ correlation of the separated twin pairs reported by Burt was strikingly high, more so than that reported in the three other studies of separated twins. The most important aspect of Burt's study, however, was that he alone had been able to measure quantitatively the similarity of the environments in which the separated twin pairs had been reared. The incredible (and convenient) result reported by Burt was that there was no correlation at all between the environments of the separated pairs.

Further, in order to fit a genetic model to IQ data, it is necessary to know what the IQ correlations are for a considerable number of types of relatives—some close and some not so close. Burt was the only investigator in history who claimed to have administered the same IQ test in the same population, to the full
gamut of biological relatives of all degrees of closeness. In fact, for some types of relatives (grandparent-grandchild, uncle-nephew, second cousin pairs), the IQ correlations reported by Burt are the only such correlations ever to have been reported. The Burt correlations for all types of relatives corresponded with remarkable precision to the values expected if IQ were almost entirely determined by the genes.

The blunt fact is that Burt's data, which have played so important a role, were reported and published in what is clearly a truly scandalous and suspicious fashion. The implausibility of Burt's claims should have been noted at once by any reasonably alert and conscientious scientific reader. To begin with, Burt never provided even the most elementary description of how, when, or where his "data" had been collected. The normal canons of scientific reporting were ignored entirely by Burt, and by the editors of the journals that published his papers. He never even identified the "IQ test" he supposedly administered to untold thousands of pairs of relatives. Within many of his papers, even the sizes of his supposed samples of relatives were not reported. The correlations were given without any supporting details. The 1943 paper that first reported many of the correlations between relatives made only the following reference to procedural details: "Some of the inquiries have been published in LCC reports or elsewhere; but the majority remain buried in typed memoranda or degree theses." Conscientious scientists usually do not refer interested readers to their primary sources and documentation in such a cavalier way. The reader should not be surprised by the fact that none of the London County Council reports, typed memoranda, or degree theses glancedly referred to by Burt have ever come to light.

The very few occasions when Burt made specific statements about his procedure should have provoked some doubts in his scientific readers. For example, in a 1955 paper Burt described the procedure by which he obtained IQ test results for parent-child, grandparent-grandchild, uncle-nephew, etc. The IQ data for children were supposedly obtained by revising (on the basis of teacher's comments) the results of unspecified IQ tests given in school. But how did Burt obtain "IQs" for adults? He wrote: "For the assessments of the parents we relied chiefly on personal interviews; but in doubtful or borderline cases an open or a camouflaged test was employed." That is, in measuring the "IQs" of adults Burt did not even claim to have administered an objective, standardized IQ test. The IQ was said to have been guessed at during an interview! The spectacle of Professor Burt administering "camouflaged" IQ tests while chatting with London grandparents is the stuff of farce, not of science. The correlations reported by Burt on this claimed basis, however, were routinely presented as hard scientific truths in textbooks of psychology, of genetics, and of education. Professor Jensen referred to precisely this work as "the most satisfactory attempt" to estimate the heritability of IQ. When Burt's procedure was publicly criticized, Hans Eysenck was able to write in Burt's defense: "I could only wish that modern workers would follow his example."

The collapse of Burt's claims within the scientific community began when attention was drawn to some numerical impossibilities in Burt's published papers. For example, Burt in 1955 claimed to have studied twenty-one pairs of separated identical twins and reported that, on some unnamed group test of intelligence, their IQ correlation was .771. By 1958 the number of pairs had been increased to "over 30"; surprisingly, the IQ correlation remained precisely .771. By 1966, when the sample size had been increased to fifty-three pairs, the correlation was still exactly .771! This remarkable tendency for IQ correlations to remain identical to the third decimal place was also true of Burt's studies of nonseparated identical twin pairs: as the sample size increased progressively with time, the correlation failed to change. The same identity to the third decimal place was also true of IQ correlations for other types of relatives published by Burt, as sample sizes increased (or in some cases decreased) over time. These and other characteristics indicated that, at the very least, Burt's data and claimed results could not be taken seriously. As one of us in 1974 concluded after surveying Burt's work: "The numbers left behind by Professor Burt are simply not worthy of our current scientific attention." The scientific exposure of Burt prompted Professor Jensen to execute a brisk about-face. Two years earlier Jensen had described Burt as a born nobleman, whose large and representative samples had secured his place in the history of science. But in 1974 Jensen wrote, after citing the absurdities that critics had already documented, that Burt's correlations were "useless for hypothesis testing"—that is to say, worthless. But Jensen maintained that Burt's work had merely been careless, not fraudulent; and he also maintained that the elimination of Burt's data did not substantially reduce the weight of the evidence demonstrating a high heritability of IQ. That incredible claim was made despite Jensen's earlier assertion that Burt's was "the most satisfactory attempt" to calculate the heritability of IQ.

The argument over Burt's data might have remained a discreet academic affair and might have tiptoed around the question of Burt's fraudulence were it not for the medical correspondent of the London Sunday Times, Oliver Gillie. Gillie tried to locate two of Burt's research associates, the Misses Conway and Howard, who had supposedly published papers in a psychological journal edited by Burt. According to Burt, they were responsible for the IQ testing of the separated identical twins, for the testing of other types of relatives, and for much of Burt's published data analyses. But Gillie could uncover absolutely no documentary record of the existence of these research associates. They had not been seen by, and were wholly unknown to, Burt's closest co-workers. When asked about them by his housekeeper, Burt had replied that they had emigrated to Australia or New Zealand, this at a time before, according to Burt's published papers, they were testing twins in England. Burt's secretary indicated that Burt had sometimes written papers signed by either Conway or Howard. These facts led Gillie to suggest, in a front-page article in 1976, that Conway and Howard may never have existed. The article flatly accused Burt of perpetrating a major scientific fraud, a charge subsequently supported by two of Burt's former students, now themselves prominent psychometricians, Alan and Ann Clarke.
IQ: THE RANK ORDERING OF THE WORLD / 159

The public exposure of Burt’s fraudulence seemed to strike a raw hereditary nerve. Professor Jensen wrote that the attack on Burt was designed “to wholly discredit the large body of research on the genetics of human mental abilities. The desperate scratch-the-earth style of criticism we have come to know in this debate has finally gone the limit, with charges of ‘fraud’ and ‘fakery’ now that Burt is no longer here to . . . take warranted legal action against such unfounded defamation.” Professor Eysenck joined in by pointing out that Burt had been “knightsed for his services” and that the charges against him contained “a whiff of McCarthyism, of notorious smear campaigns, and of what used to be known as character assassination.”

The attempt to defend Burt by assaulting his critics soon collapsed. The eulogy at Burt’s memorial service had been delivered by an admiral, Professor Leslie Hearnshaw, and had prompted Burt’s sister, in 1971, to commission Hearnshaw to write a biography of her distinguished brother and to make Burt’s private papers and diaries freely available to him. When the fraud charges exploded, Hearnshaw wrote to the Bulletin of the British Psychological Society, indicating that he would assess all the available evidence and warning that the charges of Burt’s critics could not be lightly dismissed. This warning seems to have muddled the tone of Burt’s more militant hereditary defenders. Thus, by 1978, Eysenck wrote of Burt: “On at least one occasion he invented, for the purpose of quoting it in one of his articles, a thesis by one of his students never in fact written; at the time I interpreted this as a sign of greaterness.”

The Hearnshaw biography, published in 1979, has put to rest any lingering doubts about Burt’s wholesale faking. The painstaking searches and inquiries made by Hearnshaw failed to unearth any substantial traces of Miss Conway, or Miss Howard, or of any separated twins. There were many instances of dishonesty, of evasion, and of contradiction in Burt’s written replies to correspondents who had inquired about his data. The evidence made clear that Burt had collected no data at all during the last thirty years of his life, when, supposedly, most of the separated twins had been studied. With painful reluctance, Hearnshaw found himself forced to conclude that the charges made by Burt’s critics were “in their essentials valid.” The evidence demonstrated that Burt had “fabricated figures” and had “falsified.” There is now no doubt whatever that all of Burt’s “data” on the heritability of IQ must be discarded. The loss of these incredibly clear-cut “data” has been devastating to the claim that a substantial IQ heritability was demonstrated.

But what are we to make of the additional fact that Burt’s transparently fraudulent data were accepted for so long, and so uncritically, by the “experts” in the field? Perhaps the clearest moral to be drawn from the Burt affair was spelled out by N. J. Mackintosh in his review of the Hearnshaw biography in the British Journal of Psychology.

Ignoring the question of fraud, the fact of the matter is that the crucial evidence that his data on IQ are scientifically unacceptable does not depend on any examination of Burt’s diaries or correspondence. It is to be found in the data themselves. The evidence was there . . . in 1961. It was, indeed, clear to anyone with eyes to see in 1958. But it was not seen until 1972, when Kamin first pointed to Burt’s totally inadequate reporting of his data and to the impossible consistencies in his correlation coefficients. Until then the data were cited, with respect bordering on reverence, as the most telling proof of the heritability of IQ. It is a sorry comment on the wider scientific community that “numbers . . . simply not worthy of our current scientific attention” . . . should have entered nearly every psychological textbook.

We do not view the uncritical acceptance of Burt’s data as an unusual or inexplicable “sorry comment on the wider scientific community.” The fraud perpetrated, and unwittingly propagated by the scientific community, served important social purposes. Professor Hearnshaw’s biography essentially saves the face of psychometry by probing the individual psychology of Burt to ask why he should have been moved to such fraudulence. Burt, no longer a nobleman but now victim of a debilitating and psychiatrically distressing disorder, has become the bad apple of psychometry. By 1980, when the British Psychological Society was prepared to draw up its “Balance Sheet on Burt,” there had been a closing of the ranks; the psychometric criers reiterated their belief that, despite the eviction of Burt, the residual evidence for the heritability of intelligence was strong. The social function of IQ ideology was still dominant.

NOTES

18. L. Kamin, "Heredity, Intelligence, Politics and Psychology," unpublished presidential address to meeting of the Eastern Psychological Association (1972).