Twin studies are often trotted out to “prove” the genetic basis of this or that condition (such as alcoholism, or schizophrenia). It’s assumed that if a condition co-occurs at high rate in identical twins relative to its co-occurrence rate in fraternal twins or non-twin siblings, the condition in question must be genetic.

It makes sense. You’d think nothing could be more of an airtight proof of something’s genetic basis than a study involving twins—individuals with identical genes. That’s the ultimate proof, isn’t it?

Not really.

Far from being the Gold Standard of research, most twin studies are classical examples of bad science—how not to do science. Almost all such studies start with an agenda of some kind (some of the most famous twins studies were conducted by self-avowed eugenicists); a study is invariably designed to buttress a preexisting pet theory (regardless of whether that theory is right or wrong), a la Nazi science, Soviet science, etc. (Researcher bias is a problem in all science, of course, but in twin studies it’s particularly profligate.) Lack of blinding is also a staple of twin studies, and in early studies of schizophrenia in twins, diagnostic standards were exceptionally poor (with researchers getting to decide, subjectively, on the flimsiest of grounds, who is “feeble-minded” and who isn’t). In fact, in some studies it’s not even certain that “identical” twins were actually properly identified. (Misidentification of dizygotic twins as monozygotic, and vice versa, was actually fairly common until recently.)

Many of the early twin studies are based on absurdly small sample sizes—study populations so small, so utterly without statistical validity,
that in any other context they’d be laughed off as inconsequential; yet because it’s a twin study, the data are accepted uncritically. But even if we overlook statistical-power issues, we’re left with an even bigger problem. The basic assumption of all of these studies is that if monozygous (identical) twins show high concordance rates for a particular study property (such as schizophrenia, alcoholism, left-handedness, ability to eat French fries, or whatever) relative to the rate for dizygotic (non-identical) twins, the only possible reason for the difference has to lie in the genes. In other words, it’s assumed (improperly; indeed, ludicrously) that all other variables are controlled for. It’s assumed that the only thing that could possibly account for the outcome is genetics. But when you think about it, that’s quite obviously a fallacy.

Twins aren’t raised “the same as other children.” Far from it. They’re celebrated as the wonderfully unique biological exceptions that they are. What this means in practice is that identical twins actually have much greater exposure to concordant-expectation-based learned behaviors than other children. Even if parents don’t always reinforce sameness expectations, twins themselves learn to appreciate and leverage their samenesses (in expected and unexpected ways). Are we to be surprised when a study finds concordances in this or that attribute? No kidding? Of course there are concordances! How could there not be?

Psychologist Jay Joseph explains it this way:

The mainstay of twin research is the “classical twin method,” developed in the 1920s, in which researchers compare the trait resemblance of reared together identical twin pairs (MZ, monozygotic) versus reared together same-sex fraternal (DZ, dizygotic) twin pairs. If identical pairs (who share a 100% genetic similarity) resemble each other more than fraternal pairs (who share an average 50% genetic similarity) for the disorder or trait in question, twin researchers conclude that the disorder or trait has an underlying genetic component. They arrive at this conclusion on the basis of several theoretical assumptions about twins, the most important and controversial of which is the assumption that identical and same-sex fraternal twin pairs grow up experiencing roughly
equal environments. This is known as the “equal-environment assumption” (EEA) of the twin method. Psychiatric twin studies are based on twin pairs reared together in the same home, and in most cases identical pairs resemble each other more for psychiatric disorders and behavioral traits than do fraternal pairs. Although the authors of psychiatry and psychology textbooks and other mainstream publications usually endorse genetic interpretations of psychiatric twin studies uncritically, there is a fatal flaw underlying these studies: identical twin pairs grow up experiencing much more similar environments than experienced by same-sex fraternal pairs, meaning that the equal environment assumption — upon which all conclusions in favor of genetics are based — is false. Therefore, many critics have argued that it is likely that identical-fraternal comparisons capture nothing more than identical pairs’ more similar treatment, greater environmental similarity, stronger attachment and emotional bond, and greater levels of identity confusion (feeling like they are two halves of the same whole).258

In psychology, so-called studies of “twins reared apart” are frequently cited in support of major genetic influences on traits such as IQ and personality. Over the years, however, critics have pointed to many potentially important invalidating problems and biases found in these studies.259 People uncritically assume, if twins are raised apart, the only variable that hasn’t been controlled for is genetics. But that’s never the case. Not even close.

When proponents of twin-study methodology are confronted with the fallacy in their assumptions, they sometimes reply that if monozygous twins end up behaviorally or psychopathologically more similar (more


“concordant”) than fraternal twins, it’s because of inherent similarities of constitution deriving from genetics. What’s laughable about this retort is its circular logic. Twin researchers blame twins’ similarity of behavior on the very thing they are trying to prove, invoking its preexistence as proof!

Another argument twin researchers put forward in support of the “equal environment assumption” (on which all twin studies rely) is that, while identicals may grow up experiencing more similar environments than fraternals, nonetheless identical and fraternal environments differ in aspects that are relevant to the trait in question. This is known as the “trait-relevant” framing of the equal environment assumption. (For example, witnessing trauma is a trait-relevant environmental factor for PTSD.) But even when it’s shown that identical pairs inevitably experience more similar trait-relevant environments than fraternal pairs, twin researchers still argue in favor of the validity of the twin method and the equal environment assumption by reverting to a circular argument that says identical pairs naturally tend to “elicit” more similar “trait-relevant” environments on their own by virtue of their similarity. So once again, it comes down to circular logic in support of an a priori agenda.

Maybe the biggest problem with twins research is that the results just aren’t that impressive. Although early twin studies by the notorious eugenicist Franz Kallmann found concordance rates for schizophrenia in identical twins as high as 69% (which Kallmann later changed to 86% after applying spurious “age correction factors”), subsequent studies have found much lower rates, and in fact, the later the study, the lower the rate. By the late 1980s (with schizophrenia much more precisely defined, in the DSM), some studies were reporting pairwise concordance rates of under 20% for schizophrenia in twins. The largest such study found a rate of just 11%. If monozygous twins were truly identical genetically, and if schizophrenia were truly genetic, the occurrence of schizophrenia in one twin should mandate its occurrence in the other twin at a rate approaching 100%, not 11%. Yet somehow researchers focus narrowly on the 11%.

number, rather than feeling a need to explain the missing 89%.

We now have much better, more powerful ways of probing genetics than merely looking at small populations of twins, of course. We have the full arsenal of DNA sequencing tools. But when we use these tools to look for the genetic correlates of mental illness, we come up empty-handed.

When a team of 33 scientists did a deep probe of the DNA of over 800 schizophrenia patients (and over 800 control subjects), they were unable to find any gene copy variations or single nucleotide polymorphisms (SNPs) that could correlate meaningfully with schizophrenia risk. The authors of the study concluded: “These data suggest that very few schizophrenia patients share identical genomic causation.” Unwilling to abandon the genetic hypothesis, the authors speculated that “much of schizophrenia risk is due to rare, moderate-to-high penetrance [gene] variants whose population frequencies reside somewhere below the threshold of detection of genome-wide screens.” (Emphasis added.) Wouldn’t it just be simpler to say your hypothesis is wrong?

Genetic studies of things like depression, schizophrenia, and alcoholism have been, on the whole, miserable failures. In the April 2013 edition of Biological Psychiatry, Karin Hek and her 85 coauthors (yes, 85 coauthors) tell of performing a genome-wide association study involving 34,549 individuals suffering from depression. They found only one SNP (single nucleotide polymorphism) that was “suggestive” of an association


with depression—and that particular SNP didn’t map to a known gene! (It was in a region of “junk DNA.”) With a certain degree of resignation, the authors concluded: “The results suggest that only a large sample comprising more than 50,000 subjects may be sufficiently powered to detect genes for depressive symptoms.” In other words, if we keep increasing the size of the study population, maybe we’ll find something.

But what if increasing the study size isn’t the answer?

Studies going back almost a century have determined that body height is 80% to 90% heritable; no one seriously questions this fact. (We can all agree that height is heritable—it “runs in families.”) However, at least three large, modern genetic studies have been done to find “height genes”; the largest involved over 180,000 study subjects (and 291 co-authors).264 In all, some 180 genetic loci were identified that play a role in determining a person’s height. But the 180 genomic features, put together, accounted for only 10% of observed variations in height. The rest appears to be environment.

Are we now supposed to enlarge our “study population” from 180,000 to several million, in order to find the genetic explanation for body height, just because we “know” one exists?

At some point, don’t we have to just admit “the data are the data”? It may sound radical to say so, but maybe it’s time, finally, to discard the deplorably bad twin studies (and the many fruitless attempts to map genetic features to mental illness) and start coming to grips with the complex, ugly, messy, unfortunate, multifactorial reality of familial behaviors that are, indeed, passed on in families but are not necessarily genetic (in any straightforward Mendelian way). Every scientist naturally wants to believe that alcoholism, schizophrenia, suicidality, major depression, problem gambling, bipolar disorder, and other conditions that we know are “passed down in families” somehow have a genetic basis, an underlying

molecular-genetic correlate, to which we can assign *causality*. I myself want to believe this. I was trained as a scientist. I know how compelling, how satisfying, how important, how downright *epic* a molecular-genetic explanation of mental illness would be. But *wanting* it so is not the same as *making* it so. It seems to me that if we’re honest with ourselves, if we just look *honestly* at the data we already have, we have to be willing to believe it’s possible there is no alcoholism gene, no schizophrenia gene, no gambling gene, no PTSD gene, no child-abuse gene—that really, what we’re looking at is not a hardware problem at all, not a “transistor-level” problem involving individual nucelotides, but a *software* problem. And guess what? *We know almost nothing about how the software works!* Why? Because we’ve spent the last five decades and umpteen billion dollars studying the *hardware*—the neurons, the neurotransmitters, the action potentials, the enzymes, the pharmacokinetics, the DNA, etc. These are important things to know about, and certainly some of our problems involve those things. But what if “alcoholism” isn’t a *gene product*? What if it’s a really messy, ugly bunch of real-world psychosocial stimuli and behavioral interactions and intergenerational patterns of maladaptive behavior that get learned and passed on? The fact that many psychological conditions can be treated successfully with talk therapy alone (e.g., some people can give up alcohol, can recover from PTSD, etc.) suggests—it *suggests*; it doesn’t *prove*—that the problem may *not really be at the hardware level*. And that’s actually a hopeful development, if true, because it’s much easier to reprogram software than to fix DNA.