Genetic Research in Psychiatry and Psychology
A critical overview
Jay Joseph

Introduction
The current consensus position in psychiatry and psychology is that, in addition to environmental factors, genes play an important role in causing psychiatric disorders and variation in “continuously distributed” psychological traits (such as personality and IQ). The fields of psychiatric genetics and behavior genetics have produced much of the research supporting this position, which consists of two broad areas of investigation. The first explores how a trait is distributed among various types of biologically related (and sometimes unrelated) people. The most common approaches have been the studies of families, of twins, and of people who have been adopted. These methods are sometimes grouped together under the heading kinship research. The second area is molecular genetic research, which attempts to pin down the actual genes that researchers believe underlie various traits and disorders.

Contrary to most academic and popular accounts of this research, in this chapter we will see that studies of families, twins, and adoptees are greatly flawed on several critical dimensions. Moreover, countless sensationalized media reports notwithstanding, researchers have failed to discover the genes they believe underlie DSM-defined psychiatric disorders, and normal variation in psychological traits. Due to space considerations, we will look only at the most frequently cited research methods.

If a trait is caused or influenced by hereditary factors, we would expect to find a greater concentration of it among members of the same family, roughly proportional to their degree of genetic relatedness. For example, first-degree biological relatives (such as the siblings, children, or parents of an individual) should manifest the trait more often, or resemble each other to a greater degree, than second-degree
biological relatives (such as an individual’s aunts, uncles, grandparents, grandchildren, nieces, nephews, or half-siblings). In addition, monozygotic (MZ, identical) twin pairs should resemble each other to a greater degree than dizygotic (DZ, fraternal) twin pairs.

A major aim of this chapter, however, is to show that although relatives frequently manifest traits in patterns predicted by genetic theories, these patterns frequently match the predictions made by theories of non-genetic causation as well. Thus, it is frequently difficult or impossible to disentangle these potential influences.

In the context of the upcoming analysis, the term environment refers to all non-genetic factors that could contribute to or cause the appearance of traits. Environmental factors and influences include parenting styles, peer groups, abuse, neglect, oppression, toxic chemicals, viruses, accidents, culture, attachment disturbance, racism, and so on. The term confound refers to unforeseen or uncontrolled-for factors that threaten the validity of conclusions researchers draw from their studies. In genetic research, potential confounds are usually environmental.

Heritability estimates (coefficients) are not evaluated in this chapter because they are misleading and widely misunderstood (Joseph, 2004, chapter 5). Genetic researchers produce such estimates (ranging from 0 to 100%) in reference to psychiatric disorders and psychological traits, and imply that as the percentage increases, the importance of genetic influences on the trait correspondingly increases. In particular, the “heritability of IQ” topic has been fiercely debated for decades. However, critics have argued, correctly in my view, that heritability estimates cannot tell us “how much” genes influence a given trait. Critical behavior geneticist Jerry Hirsch (1997, 2004) argued that a heritability estimate is not a “nature/nurture ratio” of the relative contributions of genes and environment. Other critics (e.g., Block, 1995; Chaufan, 2008; Feldman & Lewontin, 1975; Greenberg, 2005; Joseph, 2004; Lewontin, 1987; Moore, 2001; Schönemann, 1997; Stoltenberg, 1997; Wahlsten, 1990, 1994) have also detailed problems with the heritability concept. According to critical behavior geneticist Douglas Wahlsten (1990, p. 119), “The only practical application of a heritability coefficient is to predict the results of a program of selective breeding.” Moreover, heritability estimates are derived from the flawed research methods discussed in this chapter.

Another important yet rarely discussed issue in genetic research is the validity and reliability of concepts such as “schizophrenia,” “IQ,” “personality,” “bipolar disorder,” “criminality,” “ADHD,” etc. Establishing the validity of these constructs, and the ability to reliably identify and define them, is an important part of any research project. Yet the validity and reliability of psychiatric disorders and psychological traits is open to question (Boyle, 2002; Hill, 1983; Kirk & Kutchins, 1992; Kutchins & Kirk, 1997; Mensh & Mensh, 1991; Richardson, 2000). Psychologist Richard Bentall has rejected the concept of discrete mental disorders: “There is no clear boundary between mental health and mental illness. Psychological complaints exist on a continuum with normal behaviors and experiences” (Bentall, 2003, p. 143). And it is widely understood that there is no consensus definition of “intelligence” (Neisser et al., 1996).
In the following sections we will look at four major areas of genetic research: Family studies, twin studies, adoption studies, and molecular genetic studies. The upcoming analysis differs from the accounts of most textbooks and popular works, whose authors usually endorse twin and adoption studies as valid instruments for the detection of genetic factors. Furthermore, these accounts sometimes erroneously claim that researchers have already discovered genes for IQ, personality traits, and the major mental disorders. One purpose of this chapter is to encourage you to develop a healthy dose of skepticism about such claims.

Family studies constitute the first step in the process of determining whether hereditary factors underlie psychiatric disorders and psychological traits. Let us now turn to a discussion of how these studies are performed, and of how we should interpret their results.

Family Studies

Background

Traits and conditions have been known to run in families since biblical times (Weissman et al., 1986). The question of what causes this to occur has been the subject of debate ever since. These questions have come to form what is sometimes called the “nature-nurture” controversy. Why do some traits run in families? Is it the result of the common familial (nurture), social, and physical environments shared by families, by the common genes shared by family members (nature), or perhaps a combination of both? The current consensus in psychiatry and psychology holds that both genes and environment play an important role, although the authors of many authoritative works emphasize genetic factors over environmental factors.

The first ostensibly scientific attempt to determine whether traits run in families was the family pedigree study, in which researchers mapped a person’s family through several generations and noted which members were affected by the trait or disorder in question. The first few decades of the 20th century saw many publications containing pedigree charts of families manifesting traits such as “insanity,” “feeblemindedness,” “genius,” “criminality,” and pellagra. Figure 18.1 shows a five-generation 1911 family pedigree chart of the mating of a “feeble-minded” woman and an “alcoholic” man. The author, Charles B. Davenport, was a leading eugenicist in the early part of the 20th century. Davenport and other proponents of eugenics used such charts to support their argument that psychological traits and socially disapproved behaviors show “a strong hereditary bias” (Davenport, 1911, p. 83).

Moving on from family pedigree studies, the first type of systematic study of relatives was the family study method. In family (and adoption) studies, researchers identify persons manifesting the trait in question (called “cases”; psychiatric geneticists call such persons “probands”), and attempt to determine whether their biological relatives manifest the trait more often than the general population expectation. These relatives comprise the “index” group (a term also used in adoption research). In the
past few decades, researchers more often compare index relatives to a control group consisting of the biological relatives of people that do not manifest the trait. If a condition is found to aggregate in families, it is said to be familial. In the past, researchers considered the results of family studies (called "eugenical family studies" in the early 20th century) as proof positive that psychiatric disorders were caused by hereditary factors. In many cases, such interpretations were put forward in support of eugenics and compulsory eugenic sterilization programs (Black, 2003; Proctor, 1988). This is seen in the early German psychiatric genetic schizophrenia family studies of Ernst Rüdin (1916); and Franz J. Kallmann (1938). According to Kallmann (1938, p. xiv), "The principal aim of our investigations was to offer conclusive proof of the inheritance of schizophrenia and to help, in this way, to establish a dependable basis for the clinical and eugenic activities of psychiatry." (See Joseph, 2004, chapter 4 for further documentation of Kallmann’s and Rüdin’s enthusiastic support for eugenic practices.) However, contrary to the views of these early researchers, "familial" is not the same as "genetic." Unfortunately, many people view these terms as being synonymous, when in fact they are not. Moreover, some researchers, and more frequently reports in the popular media, continue to mistakenly cite family data in support of genetics (see Joseph, 2006).

The following description, analysis, and critique of family, twin, and adoption research draws on five works published since 1999 by authors who are among the world’s leading behavioral genetic and psychiatric genetic theorists and researchers.

---

Figure 18.1. Five generation pedigree of the offspring of a "feeble minded" woman and an "alcoholic" man, published in 1911.
The five publications are *Genetics of Mental Disorders*, by psychiatric geneticists Steven V. Faraone, Ming T. Tsuang, and Debby W. Tsuang (Faraone, Tsuang, & Tsuang, 1999), *Genes, Environment, and Psychopathology*, by psychiatric geneticists Kenneth S. Kendler and Carol A. Prescott (Kendler & Prescott, 2006), *Genes and Behavior*, by behavioral geneticist Michael Rutter (Rutter, 2006), *Behavioral Genetics*, the standard behavioral genetics textbook by Robert Plomin, John C. DeFries, Gerald E. McClearn, and Peter McGuffin (Fifth Edition; Plomin, DeFries, McClearn, & McGuffin, 2008), and “Genetic and Environmental Influences on Human Psychological Differences,” by behavioural geneticists Thomas J. Bouchard, Jr. and Matt McGue (Bouchard & McGue, 2003). Subsequent references to the “five publications” refer to the works of these authoritative authors/researchers.

The authors of the five publications agree that a family study, while constituting an important stage of genetic research, is unable to disentangle the potential role of genetic and environmental factors. In other words, because family members share a common environment as well as common genes, a trait “running in the family” can be completely explained by genetic or environmental factors. Family studies, therefore, do not supply evidence in support of genetics. The authors of the five publications weighed in on the issue as follows:

> Resemblance among relatives can be ascribed to shared environment (nurture) or shared genes (nature). (Kendler, 2000, p. 1149. Kendler & Prescott did not address family studies in their 2006 publication)

> Showing that a disorder runs in families does not conclusively establish that genes cause the disorder. Although family studies are indispensable for establishing the familial transmission of disorders, they cannot by themselves establish the causes of disorders. (Faraone et al., 1999, p. 21)

> Family studies differ from twin and adoptee designs in the key respect that they do not permit a clear separation of genetic and non-genetic influences. (Rutter, 2006, p. 58)

> Many behaviors “run in families,” but family resemblance can be due to either nature or nurture. (Plomin et al., 2008, p. 70)

> Family studies by themselves cannot disentangle genetic and environmental influences. (Plomin et al., 2008, p. 151)

> Correlations between biological relatives (i.e., IQ correlations between siblings or parents and offspring) reared together are etiologically ambiguous. Behavior geneticists are quick to point out that “familial does not equate to genetic.” (Bouchard & McGue, 2003, p. 5)

**How Family Studies are Performed**

Modern family studies in psychiatry employ DSM-defined diagnoses, control groups, structured diagnostic interviews, and blind diagnoses. These techniques
were largely absent in studies published before the 1970s. Researchers now compare the index group to a control group in order to rule out the possibility that faulty diagnostic methods were used, as opposed to the previous practice of comparing diagnostic rates against the general population expectation (Faraone et al., 1999). Ideally, the control cases would be similar to the index case in terms of age, sex, and other matching criteria, and would differ only in diagnostic status. Structured interviews are employed in order to standardize diagnostic procedures based on accepted criteria. For continuously distributed traits such as IQ and personality, standardized testing instruments can serve this purpose. Blind diagnostic procedures mean that the diagnosticians, raters, or scorers have no knowledge of whether family members are related to the index or control cases. Most of the earlier researchers were strong proponents of genetic theories, and it is likely that the non-blinded diagnoses they made were biased by their pre-existing views on genetics and eugenics.

Although the early studies contained the biases I have just described, their results continue to be presented to students in textbooks and in widely reproduced graphics such as Irving Gottesman’s Figure 10 on the lifetime risk of schizophrenia, from his book *Schizophrenia Genesis* (Gottesman, 1991, p. 96). The familial risk percentages in this figure appear to be heavily weighted by the poorly executed, non-blinded, large sample-size studies of a previous era. Modern studies using blind diagnoses, structured interviews, and control groups find much lower schizophrenia rates among first-degree biological relatives (Joseph, 2006; Joseph & Leo, 2006). However, in general both environmental and genetic theories of mental disorders predict the finding of higher rates of the disorder among the biological relatives of index cases versus controls, or versus the general population lifetime prevalence (approximately 0.5%–1.5% in the case of schizophrenia; APA, 2000).

A 1990 family study of ADD (attention deficit disorder; now called “ADHD”) by psychiatric genetic investigators Joseph Biederman and colleagues (Biederman, Faraone, Keenan, Knee, & Tsuang, 1990) provides an example of a modern psychiatric family study. As seen in Figure 18.2, the index cases consisted of 73 consecutively-ascertained outpatient children and adolescents diagnosed with ADD on the basis of DSM-III criteria and structured psychiatric interviews. A control group consisting of 52 children and adolescents not diagnosed with ADD was also established. The index and control cases produced 457 first-degree biological relatives (index N = 264, control N = 193). These relatives were evaluated and diagnosed by blinded raters, again using structured diagnostic interviews. The results showed an age-corrected lifetime risk for ADD (calculated as a “morbid risk” or “MR” by the investigators) of 25.1% for the index biological relatives, versus a rate of 4.6%–5.3% among the biological relatives of controls. This is a statistically significant difference (p < 0.00001). The researchers concluded that “ADD is a highly familial disorder” (p. 532), while recognizing that “familial aggregation does not necessarily imply genetic risk” (p. 531).
Bouchard and McGue (1981) reported that the pooled IQ correlations of the studies they surveyed were .45 for non-twin siblings reared together, .24 for siblings reared apart, and .385 for single parent and offspring. Personality test score correlations are somewhat lower (Plomin et al., 2008). In general, as expected by people with such diverging viewpoints as behavior geneticists, psychoanalysts, hereditarian theorists, critics of hereditarian theories, and family therapists, people sharing a common family environment resemble each other more for psychological traits that do randomly selected members of the population.

Thus, genetic researchers have turned to twin and adoption studies in an attempt to clearly separate the potential role of genes and environment. We will see, however, that this separation is far more difficult to accomplish than is currently believed.
Some ways that twins have been used for research purposes include:

- The twin method (twins reared-together)
- Studies of twins reared-apart
- The co-twin control method
- Genetic studies of the offspring of discordant monozygotic twin pairs
- Studies of discordant monozygotic twin pairs (investigating environmental differences)

After a brief description of these methods, the remaining portion of this section is devoted to an analysis of the twin method, and of twins reared-apart studies.

The Twin Method

The 1920s saw the development of the “classical twin method,” more commonly known as the “twin method.” The twin method compares the trait resemblance of reared-together MZ twin pairs (also known as monozygotic, or identical), who share 100% genetic similarity, versus the resemblance of reared-together same-sex DZ twin pairs (also known as dizygotic, or fraternal), who average a 50% genetic similarity. Twins’ trait resemblance is usually measured with concordance rates or correlations. Twin pairs are concordant when both are diagnosed with the same disorder, and discordant when only one member of the pair is diagnosed. Based on the theoretical assumption that the childhood and adult environments of both types of twins are comparable, known as the equal environment assumption (or “EEA”), twin researchers attribute to genetic factors the usual finding of a significantly greater resemblance among MZ versus same-sex DZ twin pairs. The EEA is, by far, the most controversial twin method assumption. The main theoretical assumptions of the twin method are outlined in Figure 18.3.

Twins Reared-Apart

In 1937, American researcher Horatio Newman and his colleagues (Newman, Freeman, & Holzinger, 1937) published the first systematic study of “reared-apart” twins. More recently, the Minnesota reared-apart twin studies of Thomas Bouchard and colleagues (Bouchard, Lykken, McGue, Segal, & Tellegen, 1990)
Reared-apart twin studies have assessed twin resemblance for psychological traits such as IQ and personality, but have not assessed twin concordance for psychiatric disorders. This is due to the difficulty of obtaining a large enough sample of reared-apart twins to perform such studies. The reports that have been published consist of case histories of individual twin pairs judged concordant or discordant for particular psychiatric disorders.

<table>
<thead>
<tr>
<th>Monozygotic (MZ) Twin Pairs</th>
<th>Same-Sex Dizygotic (DZ) Twin Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>![MZ Twin Pairs]</td>
<td>![DZ Twin Pairs]</td>
</tr>
<tr>
<td>Share 100% of the same genes. Reared together in the same home</td>
<td>Share on average 50% of the same genes. Reared together in the same home</td>
</tr>
</tbody>
</table>

A significantly greater concordance rate or correlation among MZ versus same-sex DZ twin pairs is attributable to genetic factors and is generalizable to the non-twin population, as long as all of the following five theoretical assumptions are true:

- **Most Controversial: The Equal Environment Assumption (EEA)**
  - MZ twin pairs and same-sex DZ twin pairs experience the same emotional and psychological bond with each other, as well as experiencing roughly the same social, treatment, and physical environments.

- **Other Assumptions**
  - There are only two types of twins, MZ and DZ
  - Investigators are able to distinguish between MZ and DZ twin pairs
  - The prevalence or distribution of the trait in question is the same among twins and non-twins (generalizability)
  - The prevalence or distribution of the trait in question is the same among individual MZ twins as a population, versus individual DZ twins as a population

Source: Adapted with revisions from Joseph, 2004, p. 22

**Figure 18.3.** The twin method and its assumptions.
Other Types of Twin Studies

Another way of studying twins is the co-twin control method. This method looks at environmental factors that might lead to different outcomes for twins. For example, researchers might wish to study the smoking habits of a pair of MZ twins, one of whom has been diagnosed with lung cancer. As a team of twin researchers observed, "Co-twins serve as exceptionally well-matched controls, removing or substantially reducing the effects of genetics, age, and race, as well as many unmeasured factors, such as pre-adult home environment, schools, religious upbringing, and so forth" (Herrell et al., 1999, p. 869). Other twin research approaches include genetic studies of the offspring of discordant MZ pairs (Fischer, 1971; Gottesman & Bertelsen, 1989; Kringlen & Cramer, 1989; for a critique of these studies, see Joseph, 2004, chapter 6; Torrey, 1990), and the study of discordant MZ pairs to assess for possible environmental differences (e.g., Mosher, Pollin, & Stabenau, 1971), or physiological differences (e.g., Torrey, Bowler, Taylor, & Gottesman, 1994) between members of the pair.

The Twin Method (Twins Reared Together)

The twin method provides the most frequently cited evidence in support of important genetic influences on psychological traits and psychiatric disorders. Indeed, genetic researchers look upon the twin method as one of the two "workhorses of human behavioural genetics" (Plomin et al., 2008, p. 38; the authors cited adoption studies as the other "workhorse"). Yet, as I will attempt to show, the twin method is no more able to disentangle potential genetic and environmental influences than is a family study. We will see, however, that although both research methods are clearly confounded by environmental influences, genetic researchers approach and interpret twin studies very differently than they approach and interpret family studies.

The twin method has been used widely to assess the role of genetic factors for IQ, personality, medical diseases, and psychiatric disorders. Researchers use correlations to measure the association (relationship) of continuously distributed traits, such as twins' scores on personality or IQ tests. A positive correlation is expressed as a coefficient ranging from 0.0 to 1.0. Concordance rates are used with "qualitative" traits such as schizophrenia, bipolar disorder (previously known as manic-depressive disorder), and autism. In these cases the researchers must determine whether the disorder is present or is not present, as opposed to continuously distributed traits, where test scores fall on a continuum.

Researchers ascertain a group of MZ and same-sex DZ pairs for their studies. The names of these twins are obtained from sources which include resident hospital populations, national or local twin registers, lists of twins consecutively
admitted to a facility, and clinical referrals. The researchers then determine zygosity, which refers to the method used to determine whether pairs are MZ or DZ. The ability to accurately make such a distinction is an essential aspect of twin research. Like family studies, most twin studies performed before the 1970s failed to make diagnoses blindly, and failed to use standard diagnostic procedures (or even adequately define the trait in question in many cases). Moreover, twin studies are subject to biases relating to the ascertainment procedures used (Rosenthal, 1962b).

If all twin method assumptions are valid (see Figure 18.3), significantly greater MZ versus same-sex DZ pair trait resemblance can be attributed to genetic factors. For personality traits measured with psychometric tests, MZs correlate at roughly .48, while DZs correlate at .23 (Bouchard, 1997a). IQ correlations are higher, with Bouchard and McGue reporting a pooled MZ correlation of .85, and a DZ pooled correlation of .58 (Bouchard & McGue, 1981; see Lerner, 1986, for a critical review of this publication). Although textbooks frequently report pooled MZ concordance for schizophrenia as 50%, it is closer to 25% in the more methodologically sound studies (Joseph, 2006; Walker, Downey, & Caspi, 1991). In any case, MZ concordance for schizophrenia is roughly 3–4 times greater than the concordance rate for same-sex DZ pairs. Many other psychiatric disorders show a similar pattern.

The results of twin studies in psychology and psychiatry are widely reported in textbooks, in popular books about genetics, and in the media. In the vast majority of cases the authors of these publications accept, with little or no criticism, twin researchers’ claims that the twin method provides unambiguous evidence in favor of genetics. We will soon see, however, that the plausibility of the equal environment assumption (EEA) is in serious doubt. If this assumption is false – and I will argue that it is false – studies utilizing the twin method may have recorded nothing more than the greater environmental similarity and psychological bond of MZ twin pairs as compared with same-sex DZ pairs.

Critics have argued since the 1920s that using the twin method to assess the role of genetics is dubious, since MZ twin pairs, in addition to being more genetically similar than DZ pairs, experience much more similar environments than DZ pairs. For example, twin researcher Harold Carter wrote in 1940 (p. 247) that “the assumption that the nurture influences are approximately equal for fraternal and identical twins... seems untenable to anyone who has had much contact with twins in their own social environment.” Carter went on to observe,

Identical twins obviously like each other better; they obviously have the same friends more often; they obviously spend more time together; and they are obviously treated by their friends, parents, teachers, and acquaintances as if they were more alike than fraternal twins are. (Carter, 1940, p. 247)

Family systems pioneer Don Jackson (1960) went further, and argued forcefully that schizophrenia twin research (and by implication twin research in psychology...
and other areas of psychiatry) may have recorded little more than the greater environmental similarity and “ego fusion” experienced by MZ versus DZ twin pairs. Jackson’s most telling point was that – among pairs with the same genetic relationship to each other – those pairs experiencing a more similar environment and a closer emotional bond were consistently more concordant for schizophrenia. This suggests that MZ-DZ concordance rate differences could be explained on environmental grounds (in addition to methodological bias). Although some twin researchers subsequently adjusted their methods on the basis of Jackson’s criticism (e.g., Gottesman & Shields, 1972), most continued to uphold the validity of the twin method. Today, Jackson’s critique is a forgotten document in the sense that the twin method is more popular than ever, even though none of his arguments against the twin method have ever been refuted (Joseph, 2001a, 2004).

Apart from the decades-old controversy over what the twin method actually measures (the EEA debate), critics of the twin method have pointed to a series of methodological problems. These include:

- The acceptance of unsupported theoretical assumptions
- The lack of an adequate and consistent definition of the trait or disorder under study
- The questionable reliability and validity of the trait or disorder under study
- The use of non-blinded diagnoses
- The use of diagnoses that were made on the basis of inadequate information
- The use of unreliable methods of zygosity determination (whether a pair is MZ or DZ)
- That hospital psychiatrists might have given MZ twins similar diagnoses because they were influenced by their knowledge of the twins’ common genetic heritage
- The unnecessary use of age-correction formulas
- The use of non-representative sample populations
- Small sample sizes
- The lack of an adequate description of the methods
- Investigator bias in favor of genetic conclusions

Still, despite these problems, there is little doubt that MZ pairs resemble each other more than same-sex DZ pairs for most behavioral traits and psychiatric disorders. However, we have seen that the decisive question is: What factors explain this difference? The answer to this question depends on the validity of the twin method’s equal environment assumption. In previous publications I have argued in detail that the EEA is not valid, and that, like family studies, the twin method is unable to disentangle the potential roles of genetic and environmental influences (see Joseph, 1998, 2000, 2002, 2004, 2006).

There is overwhelming evidence that MZ twin pairs experience much more similar environments than DZ pairs (Joseph, 2004, 2006). Perhaps more important,
MZ pairs experience a much stronger psychological bond than DZs, and more often experience what Jackson (1960) characterized as “identity confusion.” In a 1967 Norwegian twin study, for example, schizophrenia twin researcher Einar Kringlen found that 90% of his MZ twins (N = 75 pairs) had experienced “identity confusion” in childhood, whereas only 10% of his same-sex DZ pairs (N = 42 pairs) had this experience (Kringlen, 1967, p. 115). Clearly, the greater environmental similarity experienced by MZ versus DZ twin pairs could completely explain the former’s greater behavioral resemblance.

Contemporary twin researchers usually concede the point that MZ twin pairs experience more similar environments (see below), yet continue to uphold the validity of the twin method and the equal environment assumption (EEA) on the basis of two key arguments.

The first argument is that, although twin researchers recognize that MZ and DZ twin pair environments are in fact different, it is the responsibility of critics of the twin method to identify the “trait-relevant” environmental factors for which these two types of twins experience dissimilar environments. A group of prominent investigators provide an example of twin researchers placing the burden of proof onto critics: “it would seem that the burden of proof rests with critics of the twin method to demonstrate that ‘trait-relevant’ environmental factors are more similar for identical than same-sex fraternal twins” (Lyons, Kendler, Provet, & Tsuang, 1991, p. 126). Other examples of genetic researchers attempting to reverse the burden of proof from themselves to critics include Bouchard (1993b), and Faraone & Biederman (2000). By “trait relevant,” twin researchers mean aspects of the environment that have been shown to contribute to the trait in question. (For example, exposure to trauma contributes to post-traumatic stress disorder.)

However, as psychologist Scott Lilienfeld and his colleagues pointed out, in the context of separating science from pseudoscience,

a basic tenet of science is that the burden of proof always falls squarely on the claimant, not the critic ... Consequently, it is up to the proponents of these techniques to demonstrate that they work, not up to the critics of these techniques to demonstrate the converse. (Lilienfeld, Lynn, & Lohr, 2003, p. 3)

At other times, twin researchers point to a body of evidence purporting to have tested and upheld the validity of the EEA (e.g., Kendler, 1983; for a critical review of the “EEA test” literature, see Joseph, 2006; Pam, Kemker, Ross, & Golden, 1996; Richardson, 1998). Moreover, although faced with a similar problem, twin researchers do not make the “trait relevant” argument when discussing potential environmental confounds in family studies. In this case they are willing to concede that, because family members share a common environment (“trait-relevant” or not), one cannot draw valid conclusions in favor of genetic influences on the basis of the family resemblance of a trait.

The second argument twin researchers put forward in defense of the twin method is that MZ pairs tend to “create” more similar environments for themselves by
virtue of their greater genetically-caused behavioral resemblance. Therefore, according to this argument, the twin method’s validity is based on determining why – not whether – MZs experience more similar environments than DZs. In a key 1983 article on twin studies in psychiatry, Kendler based much of his argument in support of the EEA on this position: “Although the similarity in environment might make MZ twins more similar,” thereby invalidating the twin method, the genetically-caused “similarity in behavior of MZ twins might create for themselves more similar environments” (Kendler, 1983, p. 1416, italics in original). Following Kendler, ADHD twin researchers David Hay and colleagues wrote that, although MZ twins “may well be treated more similarly” than DZs, “this is far more a consequence of their genetic similarity in behaviour (and of ensuing responses by parents and others) than a cause of such similarity” (Hay, McStephen, & Levy, 2001, p. 12).

However, those who make this argument (including the authors of the five publications, see below) fail to understand that the reason MZ pairs experience more similar environments than DZ pairs, be it environmental or genetic, is irrelevant in assessing the validity of the EEA. For example, suppose that schizophrenia is caused solely by exposure to the chemical mercury. Because MZ pairs spend much more time together than DZ pairs, it is much more likely that both members of an MZ twin pair will be exposed to mercury, and subsequently be diagnosed with schizophrenia, than it is that both members of a DZ pair will be exposed and diagnosed. Let us further imagine that MZ twins are more genetically predisposed than DZs to enjoy spending time at the beach. Although MZ pairs may well show much higher concordance for skin cancer than DZs, this does not mean that skin cancer is a genetically-based disease.

On a psychological level, the theorized genetically-programmed behavioral resemblance of MZ pairs, and the “ensuing responses by parents and others,” could create more similar abusive, abandoning, or traumatizing parental behavior that could lead to higher concordance for childhood or adult disorders such as, for example, anxiety, depression, or psychosis (Bentall, 2003; Read, Fink, Rudegeair, Felitti, & Whitfield, 2008; Read, Mosher, & Bentall, 2004). In this case it is not heredity, but rather abuse, abandonment, or trauma that plays a major role in causing psychiatric disorders.

Thus, even if MZ pairs do indeed “create” more similar environments for themselves than do DZ pairs on the basis of their greater genetic similarity, it would be erroneous to conclude that higher MZ versus DZ concordance for schizophrenia, skin cancer, depression, or anxiety constitutes evidence that these conditions have a genetic basis. In the first example – regardless of why MZ pairs are together more often – higher concordance for schizophrenia among MZ pairs is caused solely their greater likelihood of being similarly exposed to mercury than DZ pairs.

Moreover, the “twins create their own environment” position illogically implies that parents are able to change their behavior on the basis of their children’s (twins’) behavior, but that children do not change their behavior on the basis of their parents’ behavior (Joseph, 1998).
Finally, proponents of the ‘‘twins create their own environment’’ position use circular reasoning, that is, they assume what they need to demonstrate. Moreover, their claim that twins’ behavioral resemblance is caused by genetics is based implicitly on the results of previous twin studies. In other words, modern twin researchers circularly rely on the twin method to validate the twin method, and in the process they circularly assume that twins’ behavioral resemblance is caused by genetics, in order to conclude that twins’ behavioral resemblance is caused by genetics.

Thus, the only relevant question in assessing the validity of the twin method and the EEA is whether – not why – MZ twin pairs experience more similar environments than DZ pairs (Joseph, 2004).

The Five Publications on the Validity of the Equal Environment Assumption

I will elaborate on the above-stated arguments against the validity of the equal environment assumption (EEA) in the context of a critical analysis of the five publications mentioned earlier (Bouchard & McGue, 2003; Faraone et al., 1999; Kendler & Prescott, 2006; Plomin et al., 2008; Rutter, 2006). All of these authors have attempted to validate the EEA and twin method, and most have themselves conducted twin research. If the EEA were valid, we would expect these experts to present a convincing argument that this indeed is the case.

Agreement on the Importance of the EEA, and Agreement That MZs Experience More Similar Environments

The authors of the five publications are in agreement that the EEA is a critical theoretical assumption of the twin method. For example, Rutter (2006, p. 41) wrote that in order to infer that genetic influences explain MZ-DZ differences,

it is necessary to rely on what has been called the ‘‘equal environments assumption’’ (EEA). In other words, one has to assume that the contrast between MZ and DZ pairs can be wholly attributable to genes because the environmental variation within MZ pairs should be much the same as within DZ pairs.

The authors also concede that the evidence shows that MZ twin pairs experience more similar environments than DZ pairs:

MZ twins are more likely than DZ twins to share friends and parental treatment in adolescence. (Bouchard & McGue, 2003, p. 9)

Several studies have found that the social environments of MZ twins are more similar than those of DZ twins. For example, habits, activities, personal preferences, parental treatment, and self-image tend to be more similar between MZ twins. Moreover, MZ twins are more likely to be dressed alike and are more likely to be confused for one another in childhood. (Faraone et al., 1999, p. 38)
Consistent with other studies, we found evidence that some aspects of the environment of members of MZ pairs are, on average, more similar than those of members of DZ pairs. (Kendler & Prescott, 2006, p. 124)

At first sight, [the EEA] seems a most implausible assumption. It is obvious, for example, that MZ twins are more likely to be dressed alike than are DZ twins. Also (for genetic reasons) MZ twins (within any pair) are more likely than DZ twins to be similar in their behavior, attitudes, and interests. It may safely be assumed that this is almost bound to lead to their choosing more similar experiences and also eliciting more similar patterns of interaction with other people. (Rutter, 2006, p. 41). [We will examine Rutter's insertion of the phrase “for genetic reasons” a bit later]

Plomin and colleagues (2008) did not state explicitly that MZ environments are more similar, but did allude to possible environmental differences between the two types of twins which they, like the others, believe are the result of genetics and therefore do not invalidate the twin method. In a stronger statement on the differing environments of MZ and DZ twin pairs, behavioral genetic twin researcher David Rowe wrote in 1994, "the question is not whether MZ twins receive more similar treatments (they do, and to claim otherwise would be foolish), but whether these treatments influence a particular trait" (Rowe, 1994, p. 45). And as early as 1979, twin researchers Sandra Scarr and Louise Carter-Saltzman (1979, p. 528) concluded, "the evidence of greater environmental similarity for MZ than DZ twins is overwhelming." Indeed, it is.

Trait Relevant Definition of the EEA

While it is clear that MZ pairs experience more similar environments than same-sex DZs, all authors continued to uphold the validity of the twin method and the EEA. Four of the five publications did so on the basis of the "trait relevant" definition of the EEA:

Twin studies of psychiatric disorders would . . . be in some trouble if MZ pairs had more similar environments than DZ pairs and if we could show that these environments altered risk for a particular psychiatric disorder. (Kendler & Prescott, 2006, p. 116, italics in original)

Twin studies may overestimate heritability if [MZ vs. DZ] differences in environmental similarity are etiologically relevant to the disorder under study. (Faraone et al., 1999, p. 38)

The EEA will not be violated if that [MZs eliciting more similar environments] is all that is occurring. That is because if the environments are being entirely driven by genes, it is reasonable to attribute the effects to genes provided, and only provided, that the environments that differ between MZ and DZ pairs do not have an effect on the trait being studied. (Rutter, 2006, pp. 41–42)
Behavioral geneticists call this assumption the “equal environmental similarity assumption,” a term that is somewhat misleading in that the issue is not whether MZ twins experience more environmental similarity than DZ twins, but rather whether they are more likely to share trait-relevant features of their environments. (Bouchard & McGue, 2003, p. 9)

It is worth noting that until the mid-1960s, virtually all twin researchers defined the equal environment assumption as the unqualified assumption that MZ and same-sex DZ twin pairs experience roughly equal environments. This is the “traditional definition” of the equal environment assumption. Until that time, most researchers either denied that MZ and DZ environments differed, or simply ignored the issue. As late as 1966, the authors of a World Health Organization (WHO) assessment of twin research concluded that “most shared post-natal experiences of MZ twins are probably not qualitatively different from those shared by DZ partners or even by sibs” (World Health Organization, 1966, p. 115). As Kendler (1983, pp. 1413–1414) described the definition of the EEA during the first four decades of the twin method’s existence, “According to the traditional view, because monozygotic and same-sex dizygotic twins share environmental factors to approximately the same extent, differences in concordance between the two twin types must be due to the influence of genetic factors.”

However, by the mid-1960s twin researchers were faced with the growing evidence of MZ-DZ environmental differences, with Jackson’s irrefutable (1960) critique, and with a good dose of common sense. Many began to realize that the 40-year-old critical theoretical assumption of the twin method, as it had been defined until then, was false. As one of many examples, in 1963 veteran Swedish psychiatric genetic twin researcher Erik Essen-Möller wrote as follows:

Quite obviously, then, the logical evidence furnished by the classical twin method is not unambiguous, as originally believed. A greater concordance in monozygotics must not invariably depend on their genetic identity, since also their environment may have been more similar (Essen-Möller, 1963, p. 69; for many more examples, see Joseph, 2004, pp. 171–175).

One could argue that twin researchers and others should have relegated the twin method to a place alongside the discarded pseudosciences of bygone eras. Or at minimum, they could have concluded that both family studies and the twin method were hopelessly confounded by environmental factors, and that results from these studies proved nothing about genetics. One highly regarded psychiatric genetic researcher did indeed appear to move toward this position near the end of his career. In 1979, David Rosenthal (1979, p. 25) concluded that both family studies and the twin method are “confounded,” and that “one can draw conclusions about them only at considerable risk.”
But the twin method lived on. What happened was that twin researchers began to subtly redefine the EEA away from the traditional definition to the new “trait-relevant” configuration. One of the first examples of this shift is found in a 1966 publication by twin researchers Irving Gottesman and James Shields, who wrote that the twin method would indeed have problems if the “environments of MZ twins are systematically more alike than those of DZ twins in features which can be shown to be of etiological significance in schizophrenia” (Gottesman & Shields, 1966, pp. 4–5, italics in original). Unfortunately, hardly anyone noticed or challenged this critical change in definition, which constituted an *ad hoc* hypothesis used to plug a gaping hole in twin method theory.

The trait relevant condition, however, means that the twin method and family studies have precisely the same problem, since both are subject to unavoidable environmental confounds. Yet twin researchers and popularizers of their work approach family studies and twin studies as if they were completely different animals. The logical fallacy is that the arguments they put forward in support of inferring a role for genetic factors from “trait-relevant” twin studies could just as easily be made in support of inferring a role for genetic factors from the creation of “trait-relevant family studies.” From the standpoint of environmental confounds, family studies and the twin method are not different animals. They are the same animal.

**EEA Test Literature**

Most of the authors argued that, although MZ twin pairs do indeed experience more similar environments than DZ pairs, the “EEA Test” research suggests that this does not constitute a major environmental bias in twin studies. This was Kendler and Prescott’s (2006) main defense of the EEA, which led them to conclude that the twin method derived heritability estimates of psychiatric disorders they presented in their book “are substantially correct” (p. 125). Further references to the EEA test literature include:

There have been various attempts to look for possible violations of the EEA with respect to twin studies of schizophrenia and other major mental disorders, with the conclusion that the EEA is not violated. (Rutter, 2006, p. 44)

The equal environments assumption has been tested in several ways and appears reasonable for most traits. (Plomin et al., 2008, p. 79)

Tests of the equal environmental similarity assumption have repeatedly shown that it is valid in most instances. (Bouchard & McGue, 2003, p. 9)

Interestingly, several EEA test studies (e.g., Borkenau, Riemann, Angleitner, & Spinath, 2002; Kaprio, Koskenvuo, & Rose, 1990; Kendler & Gardner, 1998;
LaBuda, Svikis, & Pickens, 1997; Lytton, 1977; Morris-Yates, Andrews, Howie, & Henderson, 1990; Scarr, 1968; Scarr & Carter-Saltzman, 1979) found that MZ pairs experience more similar environments than DZ pairs. The authors of these studies usually argue, however, that the greater environmental similarity of MZ pairs does not contribute to their greater behavioral resemblance, or if it does, that MZ's greater behavioral resemblance is caused by their greater genetic similarity.

This follows Kendler's (1983) position that "the behavioral similarity of monozygotic versus dizygotic twins cannot be ascribed to differences in treatment of the twins by the social environment" (p. 1416). This position is unsustainable because, among other reasons, it must generalize to mean that no one's behavior (whether twins or non-twins) is influenced by his or her social environment (Joseph, 2006). Individual twins, like individual non-twins (singletons), are human beings who receive treatment in their social, cultural, and familial environments. Yet Kendler argued that the social environments experienced by twins as individuals do not influence their behavior. It must therefore follow, for Kendler's EEA theory to hold, that no-one's behavior is influenced by his or her social, cultural, or familial environment.

Thus, the widely recognized greater environmental similarity of MZ versus DZ twin pairs invalidates the twin method on its face. The twin method, therefore, is contaminated by environmental factors regardless of what EEA-test researchers have claimed. What they must demonstrate – without qualification – is that MZ and DZ pairs experience roughly equal environments.

Twins Creating Their Own Environments

As we have seen, twin researchers have defended the twin method on the grounds that MZ twin pairs create (elicit) more similar environments for themselves on the basis of their more similar behavioral characteristics, which twin researchers attribute to their greater genetic similarity. In the past, twin researchers such as Kendler (1983), Scarr (1968), Shields, (1954), and Zerbin-Rüdin (1972) defended the twin method and the equal environment assumption on this basis. In fact, much of Kendler’s earlier defense of the EEA was based on his position that "the similar phenotypes in monozygotic twins are caused by their genetic similarity" (Kendler, 1983, p. 1414; reaffirmed as recently as Kendler, 2000), and that MZ twins create more similar environments for themselves on the basis of their greater genetic similarity.

However, we have already seen that the "twins create their own environment" argument does not hold up: (a) because even if it were true, it does little to support the EEA; (b) because it illogically implies that parents – but not twins – are able to change their behavior on the basis of others’ behavior; and (c) because twin researchers circularly assume that twins’ behavioral resemblance is genetic in order to conclude the very same thing. Thus, Kendler and other twin researchers simultaneously assume and conclude that "the similar phenotypes in monozygotic twins are caused by their genetic similarity."
Thus, twin researchers’ interpretations of MZ-DZ correlational or concordance rate differences as supporting a role for genetics is tautological. They argue, in essence, that the twin method is valid because... the twin method was previously shown to be valid. Furthermore, they seem to argue that the EEA is valid (a) if MZ and DZ pairs experience equal environments, or (b) if MZ and DZ pairs experience far different environments. This “heads I win, tails you lose” argument has little scientific validity.

It is worth noting that Kendler and Prescott (2006) appear to have abandoned the “twins create their own environments” argument, upon which Kendler had earlier (1983) placed so much importance. Others, such as Rutter, claimed that “(for genetic reasons) MZ twins (within any pair) are more likely than DZ twins to be similar in their behavior, attitudes, and interest” (2006, p. 41). But why add the phrase “for genetic reasons”? Rutter merely proclaimed this to be the case (again, implicitly basing his argument on the results of previous twin studies), and then concluded, “the EEA will not be violated if that is all that is occurring” (pp. 41–42). In fact, as we saw earlier in the hypothetical example of schizophrenia being caused by mercury, and in the real examples of skin cancer being caused by sunbathing, and depression, anxiety and psychosis being caused by trauma and abuse, the EEA will be violated even if this is occurring.

According to Plomin et al. (2008, p. 79), “some experiences may be driven genetically. Such differences between identical and fraternal twins in experience are not a violation of the equal environments assumption because the differences are not caused environmentally.” And Bouchard and colleagues had earlier argued,

Adult MZ twins... tend to remain in closer contact than DZ twins or other siblings, but we believe that this additional contact does not “cause” them to become more alike. We suggest instead, as the most plausible hypothesis, that MZs especially enjoy each other’s company because they are so [genetically] similar in personality, interests, and attitudes. (Lykken, McGue, Bouchard, & Tellegen, 1990, p. 560)

For the reasons I have already outlined, circular arguments of this type do little to uphold the validity of the equal environment assumption.

Equalizing MZ and DZ Environments Through Rhetoric

Several authors implied that there are environmental aspects of the MZ twinship that might make such pairs differ from one another. Kendler (1983) referred to this as a possible “reverse bias” in schizophrenia twin research. While biases of this type may well exist, the message the authors convey is that similarity biases and differentiating biases might cancel each other out. An earlier attempt to create such a rhetorical balance is found in a publication by Gottesman and Shields (1966, p. 55),
who implied that “the same proportion of potential schizophrenics are held back from overt illness by identifying with a normal twin as those who became ill by identifying with a normal one.”

We find similar unsubstantiated claims in three of the five publications. Apart from discussions of prenatal and obstetric factors, the authors provide no citations in support of their claims:

The greater physical and environmental similarity of MZ twins may actually lead to a decrease in behavioral similarity. (Faraone et al., 1999, p. 38)

A second possible threat to the EEA is provided by circumstances in which the experiences of MZ twins within the same pair tend to be less alike than those of DZ twins. When this is the case the violation of the EEA will lead to a misleading underestimate of genetic effects if the environmental influences have effects on the trait or disorder being studied. The main circumstance in which this could be the case concerns obstetric factors. (Rutter, 2006, p. 42; italics in original)

Prenatally, identical twins may experience greater environmental differences than fraternal twins . . . To the extent that identical twins experience less similar environments, the twin method will underestimate heritability. (Plomin et al., 2008, p. 79; italics in original)

Again, while it is possible that “reverse biases” in twin research exist, the authors provide no reason to reject the idea that obvious biases in the direction of creating more similar twin pairs are massively larger.

Global or Trait-Specific Acceptance or Rejection of the EEA?

The authors of three of the five publications maintained that the acceptance or rejection of the equal environment assumption is not a global evaluation, but that the EEA must be tested on a trait-by-trait basis:

There is no such thing as a “generic” violation of the EEA. Potential violations of the EEA must be evaluated disorder by disorder. (Kendler & Prescott, 2006, p. 117)

Whether or not the EEA is, or is not, violated will vary by traits. There cannot be any general conclusions on the EEA. (Rutter, 2006, p. 43)

Good scientific practice . . . requires that the [EEA] be repeatedly tested for each trait under investigation and particular findings that depend on the assumption be replicated in designs that do not make the assumption. (Bouchard & McGue, 2003, p. 9)

Once again, the authors imply a qualitative distinction between family studies and twin studies, when no such distinction is warranted. I am unaware of any
behavior genetic or psychiatric genetic researcher arguing in support of testing for environmental confounds in family studies, or that family studies should be tested “disorder by disorder” for such confounds. On the contrary, we have seen that a simple understanding that families experience common environmental influences is sufficient for these researchers to indeed reach the “general conclusion” that family studies are hopelessly confounded by environmental factors. Their argument in support of the twin method appears to be based on the logical fallacy of “special pleading,” which refers to the application of standards, principles, and rules to others while claiming to be exempt, without providing adequate justification for the exemption.

Do Other Types of Studies Validate the Twin Method?

Faraone and colleagues emphasized several times that twin method results, by themselves, prove little about genetics: “Any conclusion about the role of genes and environment must rely not on a single study or class of study but on the converging evidence provided by a variety of research paradigms” (Faraone et al., 1999, p. 43; the entire sentence was italicized in the original). Elsewhere they wrote that a “key point” in psychiatric genetic research is, “No one study proves or disproves anything. Scientists require a pattern of converging evidence from multiple studies before they can reasonably conclude that genes play a role in causing the disease” (Faraone et al., 1999, p. 12).

According to Bouchard and McGue (2003, p. 10), “inferences about the nature and existence of genetic and environmental influences on individual differences in behavior do not rest solely with twin studies.” Rutter, while recognizing the “limitations” of twin research (2006, p. 59), argued that the “overall pattern” of behavior genetic findings “demands the acceptance of the importance of genetic influences” (p. 60). These researchers echoed one of Gottesman’s earlier positions, where he wrote that although schizophrenia family, twin, and adoption studies each “contribute to the genetic argument... No one method alone yields conclusive proof or disproof” (Gottesman, 1991, p. 93).

Lilienfeld et al. (2003) addressed this “holism” argument, pointing out that pseudoscience proponents “typically maintain that scientific claims can be evaluated only within the context of broader claims and therefore cannot be judged in isolation” (p. 9). An example they gave was the response of proponents of the Rorschach Inkblot Test to their critics. Supporters of the Rorschach sometimes caution that its results should not be interpreted in isolation, but instead should be considered along with other information obtained in a psychological evaluation. For Lilienfeld et al., this means that “proponents of the Rorschach and other techniques can readily avoid subjecting their claims to the risk of falsification.” This allows them to protect their techniques through the “heads I win, tails you lose”
(p. 9) position, whereby proponents can point to positive research in support of the technique, while dismissing negative findings on the grounds that the technique should never be judged in isolation.

Clearly, while recognizing the “limitations” of the twin method, the twin researchers cited above have helped immunize the method from criticism by attempting to validate it through the results provided by the “converging evidence” allegedly supplied by other types of studies. Their error lies in this: If the EEA and other assumptions were truly sound, the twin method – standing alone – would indeed provide conclusive evidence in favor of genetics.

Would the Falseness of the EEA Invalidate the Twin Method, or Merely Lead to an “Overestimation of the Genetic Effects?”

In a 1993 article on psychiatric twin studies, Kendler wrote, “The EEA is crucial because, if the EEA is incorrect, excess resemblance of MZ twins compared with DZ twins ascribed to genetic factors could be partly or entirely due to environmental effects” (Kendler, 1993, p. 906, italics added). Kendler’s accurate assessment summarizes the point that critics of the twin method have made for over three generations. The main difference between Kendler and the critics is that Kendler affirms the validity of the EEA, whereas critics have argued the opposite position.

By 2006, however, Kendler would allow only that “Failure of the assumptions of the twin model can lead to incorrect genetic and environmental estimates” (Kendler & Prescott, 2006, p. 114). The authors of three of the four other publications made similar arguments:

MZ twins are more likely to be dressed alike and are more likely to be confused for one another in childhood. Thus, twin studies may overestimate heritability if these differences in environmental similarity are etiologically relevant to the disorder under study. (Faraone et al., 1999, p. 38)

With respect to these traits, some of the difference in similarity between MZ and DZ pairs will be due to environmental influences. This means that, to a degree, [the] EEA is violated . . . it will mean that the standard way of measuring heritability will tend to overestimate the genetic effect. (Rutter, 2006, p. 43)

If the assumption [EEA] were violated because identical twins experience more similar environmental than fraternal twins, this violation would inflate estimates of genetic influence. (Plomin et al., 2008, p. 79)

It appears that twin researchers have further immunized the twin method from falsification in the sense that, as opposed to invalidating genetic interpretations of twin method results in general, they have transformed a finding that the EEA is invalid to indicate only a degree of error that leads to an overestimation – rather than a negation – of genetic effects.
The Failure to Address the Specific Arguments of EEA Critics

Another problem found in the five publications, and indeed in the behavioral genetic and psychiatric genetic literature in general, is the frequent failure to address the specific arguments of critics of the twin method. Although genetically-oriented authors sometimes mention unnamed “critics,” they rarely quote them or take their specific objections seriously. In their discussions of twin method theory and practice, Kendler and Prescott, Faraone et al., and Plomin et al. did not even mention that critics exist, even though there have been many such critics (for example, Bleuler, 1978; Boyle, 1990, 2002; Breggin, 1991; Charney, 2008; Hoffman, 1985, 1991; Jackson, 1960; Joseph, 1998, 2000, 2002, 2004, 2006; Kamin, 1974; Kamin, in Eysenck vs. Kamin, 1981; Kamin, 1981; Lewontin, Rose, & Kamin, 1984; Marshall, 1990; Neel & Schull, 1954; Pam, 1995; Pam et al., 1996; Phillips, 1993; Richardson, 1998; Richardson & Norgate, 2005; Rosenthal, 1960, 1961, 1962a, 1962b). Only Rutter (2006) referenced a critic who pointed to post-natal environmental influences that might lead MZ twins to resemble each other more than DZs.

A Failure to Recognize the Unique Psychological Bond of MZ Twin Pairs

Twin researchers assess the childhood environments of twins by asking questions such as “Did you share the same room,” “Did you dress alike,” and “Did you and your twin have the same friends” (Kendler & Prescott, 2006, p. 117). What they tend to overlook is that the MZ twinship is a unique human relationship, which involves an extreme level of closeness, mutual association, and difficulty in maintaining a separate identity from one’s co-twin. “Identity formation,” wrote Ricardo Ainslie (1985, p. 50) in his book on twinship, “is often considered the cornerstone of any discussion of the psychology of twinship. The idea that twins encounter difficulties in the process of identity formation is as pervasive in scientific writings on twinship as it is in popular culture.”

Although the problem of identity formation might not greatly impact twins’ correlations for traits such as IQ, we could expect it to have a major impact on twin personality correlations, and on twin concordance rates for psychotic disorders such as schizophrenia (Jackson, 1960; Kringlen, 1967).

Conclusion

The twin method has supplied the most frequently cited evidence in support of important genetic influences on psychiatric disorders and psychological trait variation. The results of these studies have been put forward, largely uncritically, in countless textbooks and popular works, in review articles, and in the media.

However, the evidence suggests that, like family studies, the twin method is unable to disentangle the possible roles of genes and environment. There is, in fact,
little reason to accept that the twin method has measured anything other than the more similar treatment, greater environmental similarity, and closer psychological association experienced by MZ versus same-sex DZ pairs. Qualifications regarding “trait-relevance” or twins “creating their own environments,” or other arguments put forward by the authors of the five publications, do little to alter this conclusion.

Studies of Reared-Apart Twins

The past few decades have seen a great deal of attention paid to studies of reared-apart twins. The intuitive appeal of these studies is understandable, since studying twins separated at birth and reared apart in different families would appear to overcome the problems of environmental confounds in the twin method. Yet, we will see that these studies are also subject to environmental confounds and other biases.

Twins reared-apart (known as “TRA”) studies compare the psychological trait resemblance of reared-apart MZ pairs (known as “MZAs”) to the resemblance of reared-together MZs (known as “MZTs”), the latter serving as a control group. Some studies have included a group of reared-apart DZ pairs (“DZAs”). TRA researchers usually conclude that, because MZA correlations are far greater than zero and are comparable to MZT correlations, their results support important genetic influences on psychological trait differences. Others have cited the results of TRA studies in support of the validity of the twin method (e.g., Alford, Funk, & Hibbing, 2005), and in support of the claim that family environment has a negligible influence on human psychological development (e.g., Harris, 1998; Rowe, 1994).

Although many anecdotal reports of individual MZA pairs have been published since the 1920s, through 1979 there had been only three systematic studies of MZAs: Newman et al. (1937), who studied 19 pairs (and used MZTs and DZTs as controls); Shields (1962), who studied 44 pairs (and used MZTs and DZTs as controls); and Juel-Nielsen (1965/1980), who studied 12 MZA pairs (and other types of twins). Newman et al. reported MZA IQ correlations of MZA .67, and .91 MZT, and personality score correlations MZA .58, and MZT .56. Shields reported IQ correlations of .77 for MZAs, and .76 for MZTs. For personality, he reported MZA and MZT “Neuroticism” correlations of .53 and .38 respectively, and .61 and .42 for “Extraversion.” Juel-Nielsen reported an MZA IQ correlation of .62, but did not calculate total sample personality test correlations. The MZA IQ correlations reported by Cyril Burt (1966) are largely discredited due to allegations of fraud or unreliability, and the data he reported are no longer accepted in the scientific literature (Bouchard & McGue, 1981; Hearnshaw, 1979; Kamin, 1974).

The Minnesota Study of Twins Reared Apart (MISTRA) was initiated in 1979 by Thomas J. Bouchard, Jr. and his colleagues at the University of Minnesota. The study concluded in 2000. Since 1988 (Tellegen et al., 1988), the MISTRA group has produced numerous publications reporting test score correlations for MZAs,
MZTs, DZAs, and other types of twins. Two other TRA studies began reporting data in the 1980s – a study from Finland (Langinvainio, Kaprio, Koskenvuo, & Lönnqvist, 1984), and the Swedish Adoption/Twin Study on Aging (SATSA; Pedersen, Plomin, McClearn, & Friberg, 1988).

According to the MISTRA researchers, in their widely cited 1990 Science publication (Bouchard et al., 1990), TRA studies “provide the simplest and most powerful method for disentangling the influence of environmental and genetic factors on human characteristics” (p. 223). The researchers reported MZA and MZT IQ correlations of .69 and .88 respectively, and personality correlations of .48 and .49 (see Figure 18.4). They concluded that, because “monozygotic twins reared apart are about as similar as are monozygotic twins reared together,” their findings supported “the strong heritability of most psychological traits” (p. 223). The MISTRA results have been popularized and promoted in many popular works (e.g., Harris, 1998; Pinker, 2002; Ridley, 2003; L. Wright, 1997; W. Wright, 1998).

<table>
<thead>
<tr>
<th>MINNESOTA MZAs</th>
<th>MZTs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IQ Tests</strong></td>
<td></td>
</tr>
<tr>
<td>WAIS–Full</td>
<td>.69</td>
</tr>
<tr>
<td>Raven, Mill Hill</td>
<td>.78</td>
</tr>
<tr>
<td><strong>Personality Tests</strong></td>
<td></td>
</tr>
<tr>
<td>MPQ</td>
<td>.50</td>
</tr>
<tr>
<td>CPI</td>
<td>.48</td>
</tr>
</tbody>
</table>

Notes: MZA = monozygotic twins reared-apart. MZT = monozygotic twins reared-together. N = number of twin pairs. R = interclass correlation. WAIS = Wechsler Adult Intelligence Scale. MPQ = Multidimensional Personality Questionnaire. CPI = California Personality Inventory. MZAs and MZTs were the only twin types reported by Bouchard et al. in this publication.

Source: Based on descriptions and results reported in Bouchard et al., 1990.

Figure 18.4. The Minnesota Study of Twins Reared Apart (MISTRA) model and selected IQ and personality test results: The 1990 Science publication.
Critique of TRA Studies

Critics have pointed to several key methodological problems with TRA studies (see Farber, 1981; Joseph, 2001b, 2004; Kamin, 1974; Kamin & Goldberger, 2002; Lewontin et al., 1984; Taylor, 1980). These include: (a) it is doubtful that most studied MZAs deserve the status of having been “reared-apart,” since most pairs had significant contact with each other for many years; (b) in several studies, there were biases favoring the recruitment of MZA pairs who resembled each other more for behavioral traits than MZA pairs as a population; (c) the Minnesota researchers failed to publish life history information for the twins under study, and then denied independent reviewers access to raw data and other unpublished information (Joseph, 2004); (d) there is controversy about whether “intelligence” and “personality” are valid and quantifiable constructs; and (e) the impact that the researchers’ bias in favor of genetic explanations may have had on their results and conclusions.

Similarity bias Kamin (1974) showed that the Newman et al. and Shields TRA (twins reared-apart) studies recruited twins on the basis of similarity and of their pre-existing knowledge of each other, which meant that these MZA samples were biased toward similarity. Susan Farber, author of the exhaustive Identical Twins Reared Apart: A Reanalysis (Farber, 1981), found that, due to ascertainment bias, “approximately 90 percent of the known cases of separated MZ twins have been studied precisely because they were so alike” (p. 36). Therefore, according to Farber, the original researchers’ conclusions that genetic factors explain MZA resemblance were based on “circular reasoning” (p. 36).

Common environmental factors shared by MZAs Although rarely mentioned in popular accounts of TRA research, there are many environmental (non-genetic) factors shared by MZA pairs that would lead them to resemble each more than pairs of randomly selected members of the world’s population. We would expect a sample of the latter, of course, to correlate near zero for psychological traits on either environmental or genetic grounds. Non-genetic (in part non-familial) influences experienced by MZA pairs are seen in Box 18.1.

Thus, the fact that MZA pairs share the common factors seen in Box 18.1 would lead to us to expect them to correlate well above zero for most psychological traits. As MISTRA researchers McGue and Bouchard have recognized, age and sex effects alone can have a “substantial” impact on MZA IQ and personality correlations (McGue & Bouchard, 1984, p. 325).

The myth of the separated twins Following up on the final bulleted point in Box 18.1, it is a myth that TRA researchers studied twins whom we can legitimately regard as having been “reared-apart” (Joseph, 2004; Taylor, 1980). In her analysis of the TRA
literature published until 1980, Farber found that only 3 of the 121 reported MZA pairs were separated during the first year of life, were reared with no knowledge that they had a twin, and were studied at the time of their first meeting. “Of the 121 cases reported in the last fifty years,” she wrote, “only three are ‘twins reared apart’ in the classical sense” (Farber, 1981, p. 60). The “most accurate description of this sample,” she concluded, is “MZ twins partially reared apart” (p. 273, italics added). And there is every reason to believe that the “separation” of Bouchard and colleagues’ Minnesota MZA pairs followed a similar pattern (Joseph, 2004).

In the Swedish study (SATSA), the investigators classified twin pairs as having been “reared apart” if they had been separated before age 11 (Pedersen, Plomin, Nesselroade, & McClearn, 1992, p. 347). The average SATSA age at separation was 2.8 years (Pedersen, McClearn, Plomin, & Nesselroade, 1992, p. 257), and about 75% had some degree of contact after separation. According to the SATSA data, MZAs (average age: 65.6 years) were “separated” for an average of only 10.9 years at the time of testing (Pedersen, Plomin, et al., 1992, p. 347; at least one pair had less than one year of separation). Twins supplied information by mail, and many were not investigated personally. In the Finnish study, 12 of the 30 MZA pairs were separated after the age of five (the mean age at separation for the entire MZA sample was 4.3 years), and the degree of post-separation contact was not stated (Langinvainio, Koskenvuo, Kaprio, & Sistonen, 1984).

**Box 18.1.** Environmental (in part non-familial) influences shared by reared-apart monozygotic twin pairs (MZAs).

- They are exactly the same age (birth cohort)
- They are the same sex
- They are almost always the same ethnicity
- Their appearance is strikingly similar (which will elicit more similar treatment)
- They usually are raised in the same socioeconomic class
- They usually are raised in the same culture
- They shared the same prenatal environment
- Most studied pairs spent a certain amount of time together in the same family environment, were aware of each other’s existence when studied, and often had regular contact over long periods of time

Cohort effects  It is critically important to understand that the behavioral resemblance of reared-apart MZ twin pairs is influenced by cohort effects, which account for similarities in people’s behaviors and preferences that arise from the characteristics of the historical periods and cultural milieu in which they experience stages of life at the same time. In other words, we would expect two genetically-unrelated people of the same gender, who are born at the same time, to resemble each other more for psychological traits and behaviors than would two randomly selected members of the population.

Thus, for reasons having nothing to do with heredity, we should expect to find a much higher “video game playing behavior” correlation in the United States among pairs of randomly selected genetically-unrelated 15-year-old boys than we would expect to find among randomly selected pairs drawn from the entire 15–100-year-old male and female population of the United States. This example illustrates the central fallacy of TRA studies.

As Farber (1981, p. 77) astutely observed, MZAs are “not so much similar to each other as they are similar to people of their eras and SES [socioeconomic statuses].” According to behavior geneticist Richard Rose,

A colleague suggests that we cannot know [the importance of MZA resemblance] without necessary control data on similarities found in pairs of age-matched strangers . . . Were one to capitalize on cohort effects by sampling unrelated but age-matched pairs, born, say, over a half-century period, the observed similarities in interests, habits, and attitudes might, indeed, be “astonishing.” (Rose, 1982, p. 960)

In most cases, MZAs share at least seven different cultural influences with MZTs: national, regional, ethnic, religious, economic class, birth cohort, and gender cohort. As Rose suggested, comparing MZAs to “age-matched strangers” might reveal “astonishing” resemblances accounted for not by common genes, but by common cultural influences.

For example, a 1981 study on the relationship between age effects and personality, among biologically unrelated people, found that the mean correlation between age and California Psychological Inventory (CPI) scale scores was .28 across all 18 scales, with 10 scales showing a correlation of .35 or higher (Martin, Blair, Dannenmaier, Jones, & Asako, 1981; the CPI purports to measure normal aspects of personality, called “folk concepts,” and was used to assess personality in the MISTRA studies). If these findings reflect age effects in the general population, the influence of common age, which represents only one of many environmental similarities shared by MZAs (see Box 18.1), accounts for more than half of the reported MZA personality correlations.

One could therefore conclude that, in addition to methodological issues and biases, a critical common flaw in the twins reared-apart (TRA) studies published to date is that the investigators mistakenly compared reared-apart MZ twin pairs
(MZAs) to reared-together MZ pairs (MZTs), thereby failing to control for the fact that both MZAs and MZTs share several important non-familial environmental influences. (The MISTRA researchers (McGue & Bouchard, 1984) attempted to correct MZA correlations for age and sex effects, but these adjustments were inadequate (see Joseph, 2004).)

Potential testing bias Another issue is that researchers do not directly compare twins’ psychological test scores to each other. Rather, they compare each twin’s total score with a norm group established by the test developer, after which twins’ scores versus the norm group are compared to each other. Theoretically, members of a twin pair could answer very differently on individual questions, or perform differently on individual tasks, yet their total scores could appear to be “highly correlated.” If both twins endorse a similar number of items on a personality scale, it does not necessarily mean that they endorsed the same items. As personality researcher Paul Kline (1993) noted in relation to “empirically keyed” personality tests such as the MMPI and the CPI, “if two subjects have the same score on the scale, the scores are not necessarily psychologically equivalent” (p. 129). Thus, it is possible for two people (e.g., MZAs) to answer many questions on a personality test differently, yet still be recorded as having highly correlated test scores.

What a valid TRA study might look like In Figure 18.5, I have outlined what a scientifically acceptable TRA study might look like. Such a study would compare the psychological trait resemblance of a group consisting of MZAs unknown to each other and reared apart from birth, versus the resemblance of a control group consisting of MZTs, but, as Rose suggested, of biologically unrelated pairs of strangers sharing the non-genetic characteristics and influences also shared by MZAs. Thus, they should be the same age, they should be the same sex, they should share the same ethnicity, they should have been raised in comparable cultural and socioeconomic conditions, and they should be similar in appearance. Moreover, they should have no contact with each other until after the completion of evaluation and testing.

After concluding such a study, we might find that the biologically unrelated pairs correlate similarly to MZAs on psychological traits, which would suggest that MZA correlations are mainly, if not entirely, the result of environmental influences. Unfortunately, to my knowledge no study of this type has ever been attempted.

The Five Publications on TRA Studies

Bouchard and McGue (2003) did not discuss TRA study methodological issues, although Bouchard had previously defended his studies in several publications (see Bouchard, 1987, 1993a, 1997b). Faraone et al. (1999) did not address TRA studies, but Faraone and Tsuang (1995, p. 51) had written previously, “Since MZ twins
reared apart do not share a common environment, any phenotypic similarity must be due to genetic factors. We cannot invoke shared environment as a cause of phenotypic concordance.” In a similar manner, Plomin et al. (2008, p. 383) wrote, “Because MZ twins reared apart are genetically identical but do not share any environmental influences, the correlation directly estimates heritability.” We have seen, however, that MZA pairs do indeed share many important environmental influences. Behavior geneticists’ denial of obvious environmental influences in TRA studies is paralleled by their denial that the twin method in contaminated by environmental factors.

Rutter (2006, pp. 52–53) took a more critical approach toward TRA studies. He wrote that the “basic idea” of these studies “is a good one,” and that “At first sight, the inference that these similarities [of MZAs] reflect genetic influences seems incontrovertible. However, the design has rather more problems than are usually acknowledged.” He noted that it is very unusual to separate twins at birth, and that published reports indicate that “actual separation took place at a rather late age, leaving open the possibility of shared rearing influences.” Rutter made reference to
previous critics’ point about the similarity bias of the samples: “There are inevitable question marks over the influences that led separated twins to volunteer to participate in the research.” He also agreed with previous critics that MZA pairs were often placed in “somewhat similar” homes, and raised questions about the “rather limited reporting so far of quantitative findings presented in scientific papers that have been subject to peer review.” Rutter concluded that, although he believed that TRA studies lend support to genetic theories of behavioral differences, “there are too many queries for it to be reasonable to place great reliance on the findings.”

Conclusion

The results of TRA studies have been widely cited in scientific and popular works as supplying definitive evidence that personality traits and cognitive ability are strongly influenced by genetic factors. However, these studies contain important problems, which include the questionable “separation” of twins, the similarity bias of the samples, the failure to publish or share raw data and life history information for the twins under study (MISTRA), and the researchers’ bias in favor of genetic explanations. Finally, both TRA researchers and the popularizers of their work have failed to recognize that reared-apart MZ twin pairs share several important environmental influences that researchers did not control for.

Thus, for environmental reasons alone, we would expect MZA pairs to correlate well above zero on IQ and personality measures. It follows that we can draw no valid conclusions in support of genetic influences on psychological trait variation from TRA studies published to date (see Joseph 2001b, 2004). As Gottesman (1982, p. 351) concluded long ago, “After a quarter century of experience with twins reared together and twins reared apart, it is my conviction that twins reared apart are a wonderful source of hypothesis generation, but not a useful source for hypothesis testing.”

Adoption Studies

Background

The decision to perform psychiatric adoption studies was based on their authors’ conviction that both family studies and the twin method are potentially confounded by environmental factors. In the opinion of schizophrenia adoption researchers Seymour Kety, David Rosenthal, and their colleagues, the evidence from family and twin studies is “inconclusive ... in that it fails to remove the influence of certain environmental factors ... In the case of monozygotic twins it
has been pointed out that such individuals usually share a disproportionate segment of environmental and interpersonal factors in addition to their genetic identity” (Kety, Rosenthal, Wender, & Schulsinger, 1968, p. 345). They believed that adoption studies would finally provide a means of clearly disentangling environmental and genetic influences on schizophrenia. Adoption studies have also been used to assess the role of genes and environment for general intelligence (presumably measured by standardized IQ tests), and personality traits.

The underlying principle of an adoption study is its assumed ability to make a clean separation of genetic and environmental influences, since adoptees inherit the genes of their biological (birth) parents, but are reared in the environment of another (adoptive) family with whom they share no genetic relationship. After briefly describing how these studies are performed, however, I will attempt to show that adoption studies, like family and twin studies, are subject to major biases and environmental confounds.

The goal of the earliest adoption studies was to assess the role of genetic and environmental influences on IQ (e.g., Burks, 1928; Leahy, 1935; Skodak & Skeels, 1949). These were followed a few decades later by a new group of IQ studies, which included the Minnesota Adoption Study (MAS; Scarr & Weinberg, 1978), the Texas Adoption Project (TAP; Horn, Loehlin, & Willerman, 1979; Loehlin, Horn, & Willerman, 1989), and the Colorado Adoption Project (CAP; Plomin & DeFries, 1985; Plomin, Fulker, Corley, & DeFries, 1997).

In the 1960s, psychiatric genetic researchers extended the adoption study method to assess the role of genetic influences on psychiatric disorders. The first studies looked at schizophrenia (Heston, 1966; Kety et al., 1968; Kety, Rosenthal, Wender, Schulsinger, & Jacobsen, 1975; Kety et al., 1994; Rosenthal, Wender, Kety, Welner, & Schulsinger, 1971; Wender, Rosenthal, Kety, Schulsinger & Welner, 1974). Other psychiatric adoption studies include a schizophrenia investigation carried out in Finland (Tienari et al., 1987; Tienari et al., 2003), two adoption studies of bipolar disorder (Mendlewicz & Rainer, 1977; Wender et al., 1986), six ADHD studies (e.g., Cantwell, 1975; Sprich, Biederman, Crawford, Mundy, & Faraone, 2000), and several studies of criminality and antisocial behavior (e.g., Bohman, Cloninger, Sigvardsson, & von Knorring, 1982; Mednick, Gabrielli, & Hutchings, 1984).

Methods

The two most frequently used adoption study models in psychiatric genetics are the Adoptees method (Figure 18.6, Design A), and the Adoptees’ Family method (Figure 18.6, Design B). The Adoptees method (e.g., Rosenthal et al., 1971; Tienari et al., 2003) begins with parents (usually mothers) diagnosed with the disorder in question. The researchers then determine the prevalence of this disorder among their adopted-away biological offspring (index group). The prevalence of the
Italics indicate the first identified relatives (cases; called "probands" in the psychiatric genetics literature). Source: Adapted from Joseph, 2006, p.44.

Figure 18.6. Three psychiatric adoption study designs.

Notes: BPD = Bipolar Disorder. ADHD = Attention Deficit/Hyperactivity Disorder. Italics indicate the first identified relatives (cases; called “probands” in the psychiatric genetics literature). Source: Adapted from Joseph, 2006, p.44.
disorder is then compared with that of a control group consisting of the adopted-away biological offspring of parents not diagnosed with the disorder. The researchers conclude that a statistically significant higher rate of the disorder among index versus control adoptees suggests a role for genetic factors in causing the disorder.

The Adoptees’ Family method (e.g., Kety et al., 1968; Kety et al., 1975) begins with adoptees diagnosed with the disorder in question. A control group of non-diagnosed adoptees is also established. The investigators then attempt to identify and diagnose the biological and adoptive relatives in each group, and make statistical comparisons between these groups. Although Kety and colleagues (1968, 1975) based their conclusions on the rate of “schizophrenia spectrum disorders” among their index versus control biological relatives, Faraone and Tsuang wrote that, for the purposes of assessing genetic influences, researchers should compare the diagnostic rates of the biological versus adoptive relatives of the index adoptee group: “If the biologic relatives of ill adoptees have higher rates of illness than the adoptive relatives of ill adoptees, then a genetic hypothesis is supported. In contrast, if the adoptive relatives show higher rates of illness, then an environmental hypothesis gains support” (Faraone & Tsuang, 1995, p. 92)

The Adoptive Family method (Wender, Rosenthal, & Kety, 1968; see Figure 18.6, Design C) has been used mainly in ADHD research. This method suffers from its inability to make a comparison between the biological and adoptive relatives of the same index adoptee (Joseph, 2000, 2006). A less utilized technique is the Cross-fostering method (Wender, Rosenthal, Kety, Schulsinger, & Welner, 1974), which investigates the adopted-away children of biological parents not diagnosed with the disorder in question, who are raised by an adoptive parent eventually diagnosed with this disorder.

Correlational adoption studies of cognitive ability calculate IQ test score correlations between adoptees and various biological and non-biological relatives. A frequently cited comparison is the correlation of adoptees and their biological parents, versus the correlation of the same adoptees and their adoptive parents (see Figure 18.7, Design B). Other studies assess the IQ scores of adoptees raised in adoptive family environments in a higher socioeconomic bracket than that of their biological parents (Figure 18.7, Design A). The aim of these studies is to determine whether improved socioeconomic environments contribute to higher IQ scores.

In addition to calculating parent-adoptee correlations, some IQ and personality studies recorded correlations among different combinations of biological and adoptive relatives. For example, the TAP researchers (Loehlin, Horn, & Willerman, 1990) compared the personality test score correlations of adopted sibling pairs versus biological sibling pairs (Figure 18.7, Design D), and a 2004 CAP IQ study used a similar design (Petrill et al., 2004). These studies sometimes find a higher correlation among biological siblings when compared to adopted siblings. Another method compares the IQ correlations of adoptive parents and their biological
children, versus the correlation of these adoptive parents with their adopted children (see Figure 18.7, Design C).

Results

The authors of most psychiatric and psychological adoption studies concluded in favor of genetic influences on the trait under study. IQ studies have found that the correlation of adoptees and their biological (birth) parents is greater than the

Note: SES = Socioeconomic status.

Figure 18.7. Four IQ adoption study designs.
correlation of adoptees and their adoptive (rearing) parents (see Horn et al., 1979; Plomin et al., 1997; Scarr & Weinberg, 1978). Thus, the investigators concluded that genetic heritage is a more important factor in determining IQ than is the rearing environment. A study performed in France (Schiff, Duyme, Dumaret, & Tomkiewicz, 1982) found that the adopted-away children of unskilled workers, who were reared in the homes of adoptive families in the upper 13% of the socio-professional scale, scored 14 points higher on IQ tests when compared with children who were reared in the homes of their unskilled worker biological parents (Figure 18.7, Design A). This finding suggests that socioeconomic environments are an important factor in determining IQ scores.

Looking at normal variation in personality, although the authors of most studies found evidence that they believed pointed in the genetic direction, a CAP personality study found no significant correlation between the personality scores of adoptees and their biological parents (Adoptive family N = 245; Plomin, Corley, Caspi, Fulker, & DeFries, 1998). Rather than conclude that their carefully performed study found no evidence in support of genetic influences on personality, however, Plomin et al. utilized the “converging evidence” argument and linked their results to those of twin studies. This enabled these leading behavioral genetic researchers to conclude in favor of genetics: “Although several factors might contribute to the discrepancy between twin and adoption results, we suggest that nonadditive genetic influence, which can be detected by twin studies but not by adoption studies, is a likely culprit” (Plomin et al., 1998, p. 211).

Critical Issues in Adoption Research

The Danish-American schizophrenia adoption studies The Danish-American schizophrenia adoption studies were initiated in the early 1960s by American psychiatric genetic researchers and their Danish colleagues. These studies are widely seen as having definitively established schizophrenia as a genetic disorder, and helped support the view that most other psychiatric disorders have an important genetic component as well. A pair of behavior geneticists looked back on these studies in the late 1990s, writing, “When a single theory is monolithic in a field, contrary findings can break paradigms … It is just this role, we believe, that the first adoption studies of schizophrenia played in the 1960s” (Rowe & Jacobson, 1999, p. 14). However, the following problems, detailed by critics since the 1970s, are among those calling such claims into question (please refer to Figure 18.6, Designs A & B):

- The investigators decided to expand the definition of schizophrenia to include non-psychotic “schizophrenia spectrum disorders,” and they would not have found statistically significant results without such an expansion (Joseph, 2004, 2006). In fact, the Kety et al. 1968 study found zero cases of chronic schizophrenia in adoptive children with schizophrenia biological parents compared to adoptive children without schizophrenia biological parents (Joseph, 2004).
schizophrenia among the 65 identified first-degree biological relatives of adoptees diagnosed with a schizophrenia spectrum disorder, and Rosenthal et al. (1968) found that only 1 of the 76 adopted-away biological offspring of a parent diagnosed with a spectrum disorder had received a hospital diagnosis of schizophrenia.

- In Kety et al.’s famous 1968 study (see Design B), there is evidence suggesting that the researchers decided to change the design of their study after the initial relative group comparisons failed to obtain statistically significant results in the genetic direction (Joseph, 2004, pp. 220–222).

- The researchers failed to adequately define schizophrenia and “schizophrenia spectrum disorders.”

- In Rosenthal’s study (Rosenthal et al., 1971; see Design A), the researchers counted manic depression (bipolar disorder) as a “schizophrenia spectrum disorder” despite their insistence elsewhere that this diagnosis is genetically unrelated to schizophrenia. For example, Kety, Rosenthal, Wender, & Schul-singer wrote, “manic-depressive illness was never thought to be in the schizo-phrenia spectrum by us” (Kety et al., 1976, p. 417, italics added; see also Rosenthal, 1971). Without these manic-depressive subjects, Rosenthal would not have been able to claim statistically significant results in the genetic direction (see Lidz, Blatt, & Cook, 1981).

- In the Kety et al. studies using interviews to make diagnoses, there were inconsistencies in the way that the researchers decided to count and diagnose dead or unavailable relatives (Lewontin et al., 1984; Lidz & Blatt, 1983).

- The researchers failed to provide case history information on adoptees or relatives, and failed to study important environmental variables.

- As an earlier critic noted, in Kety’s Adoptees’ Family study (Design B), the “procedure of counting up all the possible relatives of each index case and pooling them as if they were independent samples . . . would allow some families to disproportionately affect the results” (Benjamin, 1976, p. 1130). Thus, the investigators’ decision to emphasize the diagnostic rate among individual relatives, as opposed to individual families, violated the assumption of independent observations underlying the statistical comparisons they used.

- In the Kety studies, the researchers decided to count first- and second-degree relatives with equal weighting.

- The researchers decided to include many late-separated and late-placed adoptees in their samples. This meant that, during sensitive developmental periods, these adoptees: (a) were reared for a certain period of time by their biological parents; (b) suffered a disruption of attachment bonds with their biological parents; and/or (c) were placed in unstable environments between separation and adoption.

- The investigators used substandard interviews to make diagnoses. In the Kety et al. studies, many of these “interviews” never took place, and were simply fabricated by the investigators on the basis of hospital records (Kendler &
Gruenberg, 1984; Lewontin et al., 1984). In the raw data Kety and colleagues called them “pseudointerviews,” but no mention of them appeared in any of the Danish-American investigators’ publications. Of the interviews that were conducted, the researchers believed that a five-minute doorstep conversation was sufficient to diagnose someone with schizophrenia (Paikin et al., 1974, pp. 308–310).

- The genetic bias of the investigators appeared to influence how they decided to count relatives, how they decided to define schizophrenia, the types of comparisons they decided to make, and the conclusions they reached (Joseph, 2004).

Selective Placement

A critical issue in psychiatric adoption research is the “no selective placement assumption.” Researchers must assume that factors relating to the adoption process (including the policies of adoption agencies) did not lead to the placement of certain groups of adoptees into environments contributing to a higher rate of the disorder in question. They must assume that children were placed into homes uncorrelated with the socioeconomic or presumed genetic status of their biological family. In many psychiatric adoption studies, however, the evidence suggests that index adoptees did experience more psychologically harmful rearing environments than those experienced by control adoptees (Joseph, 2004, 2006; Lewontin et al., 1984). This suggests that children whose biological family had a history of mental disorders were seen as inferior potential adoptees, and were therefore more likely to be placed into more chaotic and harmful adoptive families. Thus, adoption studies’ theoretical ability to disentangle genetic and environmental influences may not have occurred in these studies.

Adoptees subsequently becoming the subjects of schizophrenia adoption research were placed in the early-to-middle part of the 20th century in three countries: Denmark, the United States (Oregon), and Finland. However, all three countries had laws permitting the compulsory eugenic sterilization of people diagnosed with schizophrenia and other mental disorders (Broberg & Roll-Hansen, 1996; Joseph, 2004). These laws were passed on the basis of the widespread belief in these countries, in that era, that people diagnosed with schizophrenia or “insanity” were the dangerous carriers of “hereditary taint.”

Thus, if we look at schizophrenia adoption research in the context of the social and political environments in which it was performed, it is clear that the great majority of studied adoptees were given up for adoption in an era in which the compulsory sterilization of “schizophrenics” for eugenic purposes was permitted by law in the country or state in which their adoptions took place (Denmark, Finland, Oregon). Leaving aside all other problems, the likelihood that a violation of the “no selective placement assumption” occurred in these studies is reason enough to reject conclusions in favor of genetics.
Selective placement is also a potentially confounding factor in adoption studies of IQ, since adoption agencies often attempt to place children they perceive as “bright” (an assessment they make on the basis of the perceived intelligence of the children’s birthmother) into better adoptive homes. According to adoption researcher and behavior geneticist Harry Munsinger,

A . . . possible source of bias in adoption studies is the selective placement of adopted children in adopting homes that are similar to their biological parents’ social and educational backgrounds. “Fitting the home to the child” has been the standard practice in most adoption agencies, and this selective placement can confound genetic endowment with environmental influence to invalidate the basic logic of an adoptive study. (Munsinger, 1975, p. 627)

Naturally, agencies “fitting the home to the child” is a far cry from random placement, which a valid adoption study would seem to require. Elsewhere, Kamin concluded, “selective placement accounts for a considerable portion of the [IQ] correlation between unmarried [birth] mothers and their relinquished offspring” (Kamin, in Eysenck vs. Kamin, 1981, p. 123).

Range Restriction

Another problem adoption researchers must address is range restriction. Although IQ adoption studies frequently assess the correlation of adoptees and their adoptive parents, this correlation may have little meaning because adoptive parents represent a specially selected, and therefore non-representative, population.

Kamin used boxing to illustrate the problems posed by range restriction in adoption studies (Kamin, in Eysenck vs. Kamin, 1981). He pointed out that if boxing authorities decided to abolish weight divisions, we would observe high correlations between boxers’ weights and their won-loss records. “To avoid such a correlation,” wrote Kamin,

definite weight divisions have been established by boxing authorities. Fights can only take place between boxers of reasonably similar weight, and the correlation between weight and boxing success is consequently very low. We are suggesting that in terms of the environments provided for their children almost all adoptive parents – unlike biological parents – are in the heavyweight division. That would account for the lower parent-child IQ correlation observed in adoptive families. The correlation would presumably be much higher if parents who would provide poor environments wanted to, and were allowed to, adopt more often. (Kamin, in Eysenck vs. Kamin, 1981, p. 117)

Lacking an understanding of the sport of boxing and the purpose of weight divisions, one could erroneously conclude that there is no inherent relationship
between weight and boxing success. The same error in interpretation could be occurring in researchers’ finding that adoptees and their adoptive parents do not correlate highly on IQ tests.

Behavior geneticist Mike Stoolmiller (1998, 1999) argued forcefully that range restriction constitutes a major bias in adoption research. He showed that studied adoptive parents represent only 37% of the “environmental quality distribution of the full population of families” (Stoolmiller, 1998, p. 429). The main sources of this bias, according to Stoolmiller, are: (a) the selection of families who want to adopt a child; (b) the criteria adoption agencies use in allowing a family to adopt; and (c) the decision of adoptive families to volunteer to be part of an adoption study. He suggested that range restriction was a major confounding factor in American adoption studies of personality and IQ, such as the CAP and the TAP (e.g., Loehlin et al., 1989), and that it is also a factor influencing the MZA correlations reported in twins reared-apart (TRA) studies.

According to psychologists Ken Richardson and Sarah Norgate, “The effect of restricted socio-demographic factors in adoptive families, and their reflection in test score variances, is to reduce adoptive parents-adopted children correlations but not biological mothers-adopted children correlations” (2006, p. 327; see Figure 18.7, Design B). They further observed that IQ adoption studies’ assumption “that the adoption situation approximates a randomized-effects design” is not supported by the evidence (p. 319). They called for a “radical re-appraisal of the [genetic] interpretations and conclusions” found in IQ adoption study publications (p. 322).

Representativeness

Another critical assumption in adoption research is that samples of adoptees, biological parents, and adoptive parents are representative of their respective populations. However, this is rarely the case (see below). Thus, for example, the already greatly flawed and limited ADHD adoption studies (for a critique, see Joseph, 2000, 2006, 2009) are further flawed by the finding that adoptees are more likely than non-adoptees to receive an ADHD diagnosis (Deutsch, 1989; Deutsch et al., 1982; Tully, Iacono, & McGue, 2008).

Attachment Disturbance

Another issue in adoption research is the potential impact of attachment disturbance on the psychological well being of the adoptees under study. In fact, a team of critics preferred to designate this body of research “Studies of abandoned children” (Cassou, Schiff, & Stewart, 1980). Although attachment disturbance may not be an issue in studies that use children adopted away at birth, it becomes
another potentially confounding factor in cases where children are separated from their birthparent months or years after birth. According to Faraone et al., “If a child has lived with a parent for even a short period of time prior to adoption, the biological relationship will have been ‘contaminated’ by the environment created by the child’s biologic parents” (1999, p. 42).

Furthermore, research performed over the past two decades suggests that disturbed parent-child attachment patterns can influence brain development during critical developmental periods (Siegel, 1999, 2001; Shore, 2001). This body of research raises the possibility that there are environmentally-caused biological differences between the brains of some adoptees and the brains of securely-attached non-adoptees, which leads to even more questions about the generalizability of adoption research to the non-adoptee population.

Correlation vs. Mean Differences in IQ Test Scores

Potential environmental confounds aside, we have seen that IQ adoption studies have found that the IQs of adopted-away children correlate more with their biological parents than with their adoptive parents. On the other hand, studies using Design A in Figure 18.7 have shown that the biological children of poor or working-class parents show a substantial IQ increase when raised in the homes of families in the upper ranges of the socioeconomic scale (Scarr & Weinberg, 1976; Schiff et al., 1982). Generally speaking, genetically-oriented commentators focus on the correlational data, whereas environmentally-oriented commentators stress the large gains made by poor or working-class children adopted into professional or upper-class families.

In assessing this “correlation versus test score rise” issue, we should keep in mind that a correlation coefficient does not measure similarity, but only how traits vary together. As IQ hereditarian theorist Arthur Jensen correctly observed, the finding of higher adoptee-biological parent vs. adoptee-adoptive parent correlations “should not be misinterpreted as meaning that adopted children’s level of IQ is, on average, closer to that of their biological mothers than to that of their adoptive mothers” (Jensen, 1998, p. 339). He continued,

In assessing the malleability of IQ... one must take account of the mean difference between the biological mother and her adopted child and compare this difference with the mother-child difference in IQ for mothers of the same IQ and socioeconomic level who did not put their child up for adoption. (Jensen, 1998, p. 339)

Schiff and Lewontin (1986) presented a table of hypothetical yet “plausible” IQ data suggesting the possibility that “adopted children, even though they may correlate individually with their biological parents more than with their adoptive parents, are, in fact, more similar as a group to the adoptive parents than to their
biological ones’ (p. 179, italics in original). In Schiff and Lewontin’s example, seen in Table 18.1, the IQs of adoptees and their biological parents are perfectly correlated, whereas there is virtually no correlation between adoptees and their adoptive parents. However, adoptees as a group had the same mean IQ as adoptive parents as a group, and differed from their biological parents’ mean group IQ by seven points.

Schiff and Lewontin’s example shows that statistically significant adoptee-biological parent correlations are compatible with adoptees’ large IQ gains when reared in enriched adoptive family environments. And we have seen that there are several potentially confounding factors that call into question behavioral genetic researchers’ interpretations of these correlations. Thus, focusing on parent-child correlations at the expense of evaluating group mean IQ differences, as behavioral geneticists frequently do, can paint a misleading picture of the potential roles of genetic and environmental influences on intelligence (see also Walker & Emory, 1985).

The Authors of the Five Publications on Adoption Research

The authors of four of the five publications pointed to several problems in adoption research discussed above. Faraone et al. (1999) wrote that, although adoption studies “can disentangle genetic and environmental contributions to the familial aggregation of disorders” (p. 44), these studies “must be viewed with some caution due to potential methodologic problems that cloud their unambiguous interpretations” (p. 42). They regarded the “greatest limitation” of these studies to be “the fact that adoptees and their families are not representative of the general

<table>
<thead>
<tr>
<th>IQ scores</th>
<th>Biological parents</th>
<th>Children (adoptive)</th>
<th>Adoptive parents</th>
</tr>
</thead>
<tbody>
<tr>
<td>93</td>
<td>100</td>
<td>108</td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>101</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>102</td>
<td>106</td>
<td></td>
</tr>
<tr>
<td>96</td>
<td>103</td>
<td>112</td>
<td></td>
</tr>
<tr>
<td>97</td>
<td>104</td>
<td>101</td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>106</td>
<td>104</td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>108</td>
<td>111</td>
<td></td>
</tr>
<tr>
<td>103</td>
<td>110</td>
<td>102</td>
<td></td>
</tr>
<tr>
<td>104</td>
<td>111</td>
<td>110</td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>112</td>
<td>103</td>
<td></td>
</tr>
<tr>
<td>Mean = 98.7</td>
<td>Mean = 105.7</td>
<td>Mean = 105.7</td>
<td></td>
</tr>
</tbody>
</table>

population.” Therefore, “one cannot be sure that results from adoption studies will generalize to the broader population of non-adoptees” (p. 42). They noted further that adoptees as a population are at greater risk for being diagnosed with a psychiatric disorder than are members of the population of non-adoptees, and pointed to problems related to “contamination” (p. 43) of these studies by the late separation of adoptees from their biological parents. Similar to their evaluation of twin studies, however, Faraone and colleagues attempted to validate adoption research as constituting one component of the “converging evidence” in favor of genetics.

Rutter (2006) noted that, although behavioral geneticists frequently claim that the parent groups in their studies are comparable, “it is obvious that they are not” (p. 54). Biological (birth) mothers are often unmarried teenagers who are much more likely to have manifested “antisocial behavior,” and whose offspring received “sub-optimal obstetric care.” Conversely, Rutter observed that “adopting parents tend to differ systematically from other parents in being better educated, more socially advantaged, and in having low rates of psychopathology” (p. 54). Rutter also discussed problems such as the fact that prospective adopting parents are screened by agencies for psychopathology, that the adoptee population has a higher rate of psychopathology than the non-adoptee population, and selective placement, which critics see as a “built-in confound between genetic risk and environmental risk” (p. 55). Regarding selective placement, Rutter saw this problem as “far less important” than that posed by the “non-representativeness of adoptees,” and the “major differences between the biological and adoptive parents” (p. 56).

Plomin et al. (2008) also raised the issue of representativeness: “If biological parents, adoptive parents, or adopted children are not representative of the rest of the population, the generalizability of adoption results could be affected” (p. 76). They pointed to the potential problems of range restriction, adoptees and birth mothers sharing prenatal environments, and selective placement, which “could cloud the separation of nature and nurture by placing adopted-apart ‘genetic’ relatives into correlated environments” (p. 77).

According to Bouchard and McGue (2003), the adoptive homes into which agencies place adoptees “may be overly homogeneous,” which could lead to “an underestimation of shared environmental effects” (p. 10). They also wrote, in explicit agreement with Stoolmiller’s position on range restriction, “if adopted children were only placed in high-income families and never reared in poverty, then environmental effects associated with family income and poverty would never be revealed in an adoption study” (p. 10).

As an example of how unaccounted-for or misunderstood bias can lead to false conclusions, we turn to the National Football League (NFL). Although most games are played on Sunday, there is a game played each Monday night (Monday Night Football). Football analysts have long suggested that teams playing on Monday night are at a competitive disadvantage for the following Sunday’s game because,
compared to their upcoming opponent, they have roughly 30 fewer hours to recover from injuries and to prepare for their next opponent.

Suppose a team of researchers decides to test this “competitive disadvantage hypothesis” by examining the won-loss records of teams playing Sunday games the week after they play on Monday night. The records go back to 1970, so they would have several decades of data at hand. The researchers write that the null hypothesis states that there is no competitive disadvantage, and that Monday night teams would win 50% of their games the following Sunday. Suppose the researchers find that teams playing on Monday night have a 49% winning record on the following Sunday, a result not significantly different from 50%. They would probably conclude that playing on Monday night does not lead to a competitive disadvantage in games played the following Sunday.

But the researchers would be mistaken, because teams chosen to appear on Monday Night Football are not randomly selected from the pool of all 32 NFL teams. Rather, the NFL and the television network select teams that did well in the previous season (especially Super Bowl winners) to appear two or three times on Monday night, whereas teams that did poorly appear one time, or not at all. Because teams are not chosen randomly, and because teams with the best win-loss record appear more often, it might be more realistic for the null hypothesis to expect these teams to win 70% of their games the following Sunday. With the expectation adjusted to account for the lack of random assignment, the researchers would likely arrive at a completely different conclusion, namely, that their research suggests that playing on Monday night does indeed place teams at a competitive disadvantage (49% winning percentage vs. the expected winning percentage of 70%). (Another factor that might affect these results is that the NFL gives Super Bowl participants more difficult schedules the following season.)

Analogous types of confounding biases appear to be operating in adoption research, as most of the authors of the five publications concede. This suggests that results, appearing at first glance to be clear-cut, are not so clear-cut upon further investigation. And we have already seen that family and twin studies contain their own set of potentially confounding biases.

Conclusion

Adoption studies provide another example of an area of kinship research where, although well-performed studies might initially seem to provide a clear separation of nature and nurture (and many were not well-performed), further analysis shows that it is very difficult to achieve such separation. The authoritative authors of the five publications were able to identify several important problem areas in adoption research. At the same time, we have seen them defend genetic interpretations of the twin method, even though its ability to disentangle potential environmental and genetic factors is even more questionable than in adoption studies.
Molecular Genetic Research

It is no secret that our field has published thousands of candidate gene association studies but few replicated findings. (Psychiatric genetic researchers Faraone, Smoller, Pato, Sullivan, & Tsuang, 2008, p. 1)

Mental disorders take a staggering health and economic toll . . . Yet progress in understanding the underlying causes of these conditions seems to be moving at a crawl . . . decades of futile hunting have made it painfully clear that the contribution of any single gene to disease is probably minuscule. (Editorial in Nature, 7/10/2008; cited as Anonymous, 2008, p. 137)

Despite progress in risk gene identification for several complex diseases, few disorders have proven as resistant to robust gene finding as psychiatric illnesses. The slow rate of progress in psychiatry and behavioral sciences partly reflects a still-evolving classification system, absence of valid pathognomonic diagnostic markers, and lack of well-defined etiologic pathways. Although these disorders have long been assumed to result from some combination of genetic vulnerability and environmental exposure, direct evidence from a specific example has not been forthcoming. Few if any of the genes identified in candidate gene association studies of psychiatric disorders have withstood the test of replication. (Molecular genetic researcher Neil Risch et al., 2009, p. 2363)

No patient, not a single one, has ever benefited from genetic research into mental illness, although many have been indirectly harmed by it (because it has discouraged the development of adequate services for patients and, during one shameful period, was used to justify their slaughter). No effective treatments have so far been devised on the basis of genetic information and, given what we now know, it seems very unlikely that further research into the genetics of psychosis will lead to important therapeutic advances in the future. Indeed, from the point of view of patients, there can be few other areas of medical research that have yielded such a dismal return for effort expended. (Psychologist Richard Bentall, 2009, p. 145)

The ongoing search for the genes believed to underlie psychiatric disorders and psychological traits is based on the consensus opinion that family, twin, and adoption studies have conclusively established an important role for hereditary factors. For example, a team of schizophrenia molecular genetic researchers justified their work on the grounds that “family, twin, and adoption studies have demonstrated that schizophrenia is predominantly a genetic disorder with a high heritability” (Brzustowicz et al., 2004, p. 1057). According to Kendler (2005a, p. 1248), kinship studies have provided “convincing evidence” that “genes that affect risk for [psychiatric] disorders must exist somewhere in the human genome.” (The genome is defined as the total genetic material of an organism or species.) And Plomin and his team of IQ molecular genetic researchers wrote in 1995, “Family, twin, and adoption studies constantly converge on the conclusion that general
cognitive ability (‘g’), often indexed by scores on intelligence (IQ) tests, is one of the most highly heritable behavioral traits” (Plomin et al., 1995, p. 107).

However, the evidence from kinship studies of families, twins, and adoptees is, as we have seen, very far from convincing. In the following pages we will examine psychiatric and psychological molecular genetic research in the context of the ongoing fruitless search for genes. In the process, we will consider some possible explanations for these failures.

Research Methods

Molecular genetic investigators use several types of research methods. In a linkage study, they attempt to identify genetic markers associated with a presumed disease gene among blood relatives. Findings are often represented as a logarithm of odds (LOD) score, which expresses the probability that the linkage occurred by chance. Although valid linkage results identify areas of the genome where relevant genes might be located, they are unable to identify specific genes. A genome scan analyzes the complete genome of an individual against a set of markers whose positions on the chromosomes are known, and then looks for common patterns of inheritance between these markers and the disease characteristics. Association studies compare the frequency of genetic markers among unrelated affected individuals and a control group, and are performed with population-based case-control, or family-based samples. A genetic marker is defined as a segment of DNA with an identifiable physical location on a chromosome, whose inheritance can be followed.

A more recently developed method, upon which genetic researchers have pinned much hope, is the genome-wide association study (GWAS; Hirschhorn & Daly, 2005; Psychiatric GWAS Consortium Steering Committee, 2009). Previously, many leading psychiatric genetic researchers had pinned their hopes on the completion of the Human Genome Project (see Goldsmith, Gottesman, & Lemry, 1997; Hyman, 1999, 2000; Potash & DePaulo, 2000). Some psychiatric genetic researchers have concluded that the “linkage era” has been a failure: “The linkage era (1980–2005) for psychiatric disorders failed to identify any single locus that was unequivocally replicated across multiple independent samples” (Burmeister, McInnis, & Zöllner, 2008, p. 528). Thus, we have entered the “era of the genome-wide association study” (Maher, Riley, & Kendler, 2008, p. 1042). A description of how a GWAS is performed can be found at a US National Institutes of Health website:

To carry out a genome-wide association study, researchers use two groups of participants: people with the disease being studied and similar people without the disease. Researchers obtain DNA from each participant, usually by drawing a blood sample or by rubbing a cotton swab along the inside of the mouth to harvest cells. Each person’s complete set of DNA, or genome, is then purified from the blood or
cells, placed on tiny chips and scanned on automated laboratory machines. The machines quickly survey each participant’s genome for strategically selected markers of genetic variation, which are called single nucleotide polymorphisms, or SNPs. If certain genetic variations are found to be significantly more frequent in people with the disease compared to people without disease, the variations are said to be “associated” with the disease. The associated genetic variations can serve as powerful pointers to the region of the human genome where the disease-causing problem resides. However, the associated variants themselves may not directly cause the disease. They may just be “tagging along” with the actual causal variants. For this reason, researchers often need to take additional steps, such as sequencing DNA base pairs in that particular region of the genome, to identify the exact genetic change involved in the disease. (National Institutes of Health, 2008)

Like previous linkage and association studies, a GWAS is susceptible to the production of false positive results (Pearson & Manolio, 2008). To date, GWAS gene finding efforts in psychology and psychiatry have failed to produce consistently replicated results (Akil et al. 2010; Plomin et al., 2008; Risch et al., 2009).

Types of Theorized Genetic Transmission

Genetic researchers postulate two main types of genetic transmission for the disorders they study. The first is Mendelian inheritance, in which a disease or trait is passed from parents to offspring by a single dominant, recessive, or sex-linked gene. Medical disorders such as Huntington’s disease and PKU are caused by a person inheriting a single disease gene. Genetic researchers now believe that it is very unlikely that any psychiatric disorder is caused by a single gene.

The second type of genetic transmission is polygenic inheritance, meaning that many genes of varying effect sizes contribute to the appearance of a disorder. This means that investigators look for several genes, or individual genes thought to have a large effect size. Most researchers believe that environmental factors or triggers are necessary to bring about disorders in people presumed to be susceptible on the basis of polygenic inheritance. Psychiatric genetic and behavioral genetic researchers believe that most of the traits and disorders they study are caused by the actions of many genes (possibly hundreds) in combination with environmental factors. Researchers believe that these “multifactorial complex” traits and disorders are the result of “a complex interacting admixture of multiple genes and multiple environmental risk factors” (Rutter, 2001, p. 227).

The Fruitless Search for Genes

The work carried out by molecular genetic researchers in psychiatry and psychology is characterized by the stunning failure to identify genes, even as countless
media reports, often based on overly optimistic or premature claims by the original researchers, continue to misleadingly suggest otherwise. (Internet searches for topics such as “schizophrenia gene discovery” or “autism gene discovery” provide many such examples.) A 2006 edition of the Wall Street Journal found science writer Sharon Begley conveying the following misinformation:

As tough as neuroscientists have been on Freud – replacing his quaint notions of ego and id with neurotransmitters and brain circuits – geneticists have struck the unkindest blow, linking depression, neuroticism, impulsivity, sexual orientation and more to people’s 25,000 or so genes. The complicated tapestry of the mind woven by Freud, a respected neuroscientist in his day, has been reduced to a four-letter genetic code. (Begley, 2006)

Psychiatric genetic researchers of the 1980s believed that they would identify genes for the major psychiatric disorders by the end of that decade (McInnis & Potash, 2004; Propping, 2005). As we saw Faraone et al. acknowledge in 2008, however, “It is no secret that our field has published thousands of candidate gene association studies but few replicated findings” (Faraone, Smoller, et al., 2008, p. 1). One could go further and argue that the psychiatric genetics field has produced no consistently replicated findings (see Akil et al., 2010). And even in cases where such associations are claimed, we must keep in mind the maxim that correlation (association) does not imply cause.

Indeed, sustained worldwide research over the past few decades has failed to identify genes presumed to underlie conditions or traits such as addictions (Buckland, 2008), ADHD (Faraone, Doyle, et al., 2008; Waldman & Gizer, 2006), anxiety disorders (Smoller, Gardner-Schuster, & Covino, 2008), autism (Akil et al., 2010; Burmeister et al., 2008; Losh, Sullivan, Trembath, & Piven, 2008), bipolar disorder (Plomin et al., 2008; Craddock & Sklar, 2009), major depressive disorder (Risch et al., 2009), obsessive-compulsive disorder (Pauls, 2008), personality disorders (Reichborn-Kjennerud, 2008), and schizophrenia (Akil et al., 2010; Bergen et al., 2010).

Turning to the search for the genes believed to underlie general cognitive ability (which researchers theorize as “quantitative trait loci,” or “QTL”), Plomin et al. recognized that, after the initial failures of the mid-1990s, “Dozens of studies have subsequently explored other candidate gene associations with g [general cognitive ability] but none have shown consistent results” (Plomin et al., 2008, p. 170). Molecular genetic investigations into personality trait variation have suffered a similar fate. As Plomin and colleagues acknowledged, the “replication of [personality] associations has been difficult.” They proposed “employing powerful strategies using mouse models” (Plomin et al., 2008, p. 263).

The following is a partial list of major problem areas in psychological and psychiatric molecular genetic research (for more details, see Joseph, 2006, Chapter 11):
The field is massively plagued by false positive results (Abbott, 2008; Faraone, Smoller, et al., 2008). Clearly, some type of systematic error is common to many or most of these studies (see Ioannidis, 2005; Wacholder, Chanock, Garcia-Closas, El ghormli, & Rothman, 2004).

Researchers (usually mistakenly) assume that previous family, twin, and adoption studies have definitively established the genetic basis of the trait or disorder under study. Few have subjected this body of research to critical analysis.

Researchers frequently interpret negative results as evidence that the trait is more complex than they originally believed. Proponents of genetic theories have rhetorically transformed years, if not decades, of fruitless gene finding efforts into evidence of the “complex genetic nature” of psychiatric disorders and psychological trait variation. It seems the more failures that are recorded, the more “genetically complex” these traits and disorders become.

The validity and reliability of “continuously distributed” psychological traits such as IQ and personality, the establishment of which is a prerequisite for performing genetic research in psychology, is questionable.

The validity and reliability of psychiatric disorders, the establishment of which is a prerequisite for performing genetic research in psychiatry, is questionable (see Kirk & Kutchins, 1992).

Some methods and accompanying statistical calculations assume that some type of genetic transmission is occurring, although it is possible that, in reality, no genetic transmission is occurring.

The association (correlation) of a gene and a trait does not mean that the gene causes the trait, and a basic principle of statistics is that “correlation does not imply cause.” There are, in fact, several non-causal explanations for gene-trait correlations (see Page, Varghese, Go, Page, & Allison, 2003).

Even if a gene is a “necessary component” of a trait, it does not necessarily mean that the gene causes the trait (Ratner, 2004).

Researchers rarely consider the possibility that the genes they are searching for do not exist.

It is often assumed that the discovery of genes would be an important achievement. However, focusing research and money on environmental interventions might be a far better course, even if genes actually are involved.

The Four Stages of Molecular Genetic Research in Psychiatry and Psychology

The justification for conducting molecular genetic research in psychiatry and psychology has followed a series of stages, which are described below.

Stage one Researchers and the authors of influential secondary sources argue, usually ignoring or dismissing the publications of critics, that previous kinship research (family, twin, and in some cases adoption studies) has established the
genetic basis of the trait or disorder in question. For most traits, genetic researchers believe (a) that genes exist and await discovery, and (b) that finding genes would aid in the understanding, treatment, or prevention of the trait. Researchers then proceed to search for these presumed genes at the molecular level. They sometimes place gene-finding efforts on the same level as the search for the cure of a deadly disease, or the virus causing an epidemic. The unfounded assumption is that we cannot understand or prevent mental disorders until we know their underlying genetic structure.

Stage two This stage involves speculation about what type of genetic transmission might be occurring. Researchers usually reject single gene theories after the initial failures to find such genes, and then put forward theories about polygenic inheritance (the actions of several genes of various effect sizes). This sometimes involves further speculation about how many genes may be involved, and on which chromosomes they might be located. Meanwhile, failed gene finding efforts continue to pile up.

Stage three This has been called the rhetoric stage (Joseph, 2006). At this point molecular genetic researchers and their supporters choose not to emphasize the unexpected failure to find genes for psychological traits and psychiatric disorders. Instead, they argue that the task of finding genes for “multifactorial complex disorders” is more difficult than they first imagined. They often predict that discoveries in the 21st century “post-genomic era” are coming soon, and speculate on the direction their research will take after they find genes. For example, psychiatric geneticists Maher et al. (2008, p. 1043), instead of emphasizing decades of failure, instead emphasized optimism:

The search for the genetic causes of schizophrenia has been a focus of both psychiatry and genetics for nearly a century. Evidence is beginning to emerge for the involvement of two different underlying mechanisms: genomic and genetic variation. Although it is premature to declare a new dawn, some rays of light are providing new directions for research.

These researchers speak of “evidence . . . beginning to emerge,” and “rays of light . . . providing new directions,” instead of the more obvious conclusion that three decades of schizophrenia gene finding efforts have uncovered zero schizophrenia genes.

At other times, prominent researchers simply proclaim that genes have been discovered. An example is C. Robert Cloninger’s subsequently unsubstantiated 2002 claim, where he wrote in a leading scientific journal of a “watershed event” in psychiatry, where for “the first time, specific genes have been discovered that influence susceptibility to schizophrenia” (Cloninger, 2002, p. 13365). Four years earlier, Plomin and Rutter had written, “Genes associated with behavioral
dimensions and disorders are beginning to be identified” (Plomin & Rutter, 1998, p. 1223). And ten years before that, in their 1988 Annual Review of Psychology contribution, Loehlin, Willerman, and Horn wrote, “We are witnessing major breakthroughs in identifying genes coding for some mental disorders” (Loehlin et al., 1988, p. 124).

Researchers continue to view psychiatric disorders such as ADHD, bipolar disorder, autism, and schizophrenia as multifactorial complex disorders even after the initial gene finding failures, and then view subsequent failures as additional evidence of the “complex” nature of the “disorder.” A pair of prominent autism genetic researchers displayed such reasoning when they wrote that the “current lack of success in finding genes for autism is similar to that of complex diseases” (Volkmar & Pauls, 2003, p. 1136). In fact, the “lack of success” in finding genes is currently a defining feature of “multifactorial complex” traits and disorders in psychiatry and psychology (Joseph, 2006).

Stage four This “throw in the towel” stage (Joseph, 2006) has yet to occur in psychiatric and psychological molecular genetic research. However, fruitless gene finding efforts cannot go on forever, and, assuming that future searches continue to come up empty, funding and support will eventually run out. Of course, the failure to discover genes does not prove that they do not exist. What is striking, however, is that researchers rarely entertain the mere possibility that the genes they are searching for do not exist.

For example, in a 2010 “Policy Forum” article published in Science (Akil et al., 2010), three Nobel prize winning researchers and their colleagues, while recognizing a “frustrating lack of progress” (p. 1580) in understanding the genetics of mental disorders, asked for one billion US dollars in genome research money over ten years. They saw this as “a very small price to pay to reduce or eliminate the awful misery and burden to society caused by mental illness” (p. 1581). (Some researchers have raised the possibility that the genes they are searching for do not exist; see DeLisi, 2008; Hardy, Low, & Singleton, 2008; Sullivan et al., 2008).

An alternative explanation holds that the failure to find genes can be explained on the grounds that the basic premise of molecular genetic research – namely, the assumption that family, twin, and adoption studies have provided definitive evidence that sought-after genes actually exist – is wrong. In 2009, a leading group of psychiatric genetic researchers acknowledged, “it is unlikely but formally possible” that current interpretations of these studies “are substantially incorrect” (Psychiatric GWAS Consortium Steering Committee, 2009, p. 15). If the basic argument I have laid out in this chapter is correct, which is that these studies do in fact provide “substantially incorrect” evidence, then it is inevitable that the fields of psychology and psychiatry will be compelled to undertake a massive re-examination of the methodology and assumptions of these kinship studies. Such re-examinations are often the basis of rejecting outmoded and unsupported paradigms, and lead to the creation of new paradigms (Kuhn, 1996).
Conclusion

Media reports and some researchers’ claims notwithstanding, molecular genetic researchers have failed to uncover the genes that they believe underlie psychiatric disorders and psychological trait variation. We have seen that the field is characterized by “decades of futile hunting.” In some cases, such as cognitive ability, autism, and ADHD, the search has been going on for over a decade. In other cases, such as schizophrenia and bipolar disorder, it has been going on since the 1970s. (Over three decades ago, a pair of bipolar adoption study researchers wrote, “A genetic vulnerability to manic-depressive disorder has been demonstrated by family, twin, and linkage studies”; Mendlewicz & Rainer, 1977, p. 327.) Researchers and their backers should revisit the question of whether this research continues to be worth the cost and resources it requires, in light of the very real possibility that the genes they have been searching for do not exist.

Summary and Conclusions

In this chapter I have described and analyzed the most frequently used and cited genetic research methods in psychiatry and psychology. We have seen that family, twin, and adoption studies contain numerous flaws and questionable underlying theoretical assumptions. While most researchers view the “converging evidence” from kinship research as showing conclusively that genes play an important role, a plausible alternative hypothesis holds that these studies fail to provide scientifically acceptable evidence in support of genetic influences on psychiatric disorders and psychological trait variation. The ongoing failure to discover the genes believed to underlie these disorders and traits provides additional support for this position.

A few years ago, I published (Joseph, 2005b) a brief critique of Kendler’s (2005c) attempt to reconcile the failure of “gene finding methods” with the results of family, twin, and adoption studies, which Kendler (p. 6) viewed as demonstrating the importance of “gene risk factors . . . for nearly all psychiatric and drug abuse disorders examined to date.” In his response to my critique, Kendler wrote,

It is one thing to criticize the methodology of specific studies. It is quite another to suggest, as Dr. Joseph does, that we reject the results of an entire field of scientific inquiry. This might have been warranted for some pseudoscientific systems, such as astrology, alchemy, and the Ptolemaic astronomic system. It is highly unlikely that modern psychiatric genetics will be judged by future historians of science to be in such company. (Kendler, 2005b, p. 1986)

Although it may be premature to “reject the results of an entire field of scientific inquiry,” an extensive critical review of these results by the fields of psychiatry and
psychology is long overdue. After this review is complete, it may become more likely than Kendler believed in 2005 that future historians will view his discipline in a similar light as historians now view the pseudosciences of previous eras.

Genetic research is usually performed by people who believe strongly that genetic factors play an important role. It is reasonable to expect this bias to influence their research and their conclusions. Over a generation ago, psychologist George Albee concluded that his early belief that social scientists discover facts in order to build theories was wrong, and that:

it is more accurate to say that people, and particularly social scientists, select theories that are consistent with their personal values attitudes, and prejudices, and then go out into the world, or into the laboratory, to seek facts that validate their beliefs about the world and about human nature, neglecting or denying observations that contradict their personal prejudices. (Albee, 1982, p. 5)

The understanding that researchers’ conclusions are influenced by their belief systems, and their personal and professional interests, simply means that critical analysis of research should become the default mode in science in general, and in human genetic research in particular (Joseph & Baldwin, 2000). This implies that a shift from confirmation to falsification is in order.

But is it really possible that a scientific method supported by experts for decades could turn out to be bad science? The case of “comparative bullet-lead analysis” provides a recent example. Beginning in the early 1960s, the FBI provided expert testimony in support of this technique, which “used chemistry to link crime-scene bullets to [unused] ones possessed by suspects on the theory that each batch of lead had a unique elemental makeup” (Solomon, 2007). However, by 2004 the US National Academy of Sciences concluded that comparative bullet-lead analysis is “unreliable and potentially misleading” (Solomon, 2007), and the method was no longer admissible as evidence. Perhaps a thorough investigation into the claim that twin and adoption studies provide conclusive evidence in favor of genetics would also find such claims to be “unreliable and potentially misleading.”

David Rosenthal (1968, p. 414) once astutely observed that “hereditarians” like to look at “numbers,” whereas “environmentalists” like to look at “patients” (people). The behavioral and psychiatric genetic literature tends to emphasize numbers and statistics at the expense of understanding the complexity of people’s histories and life circumstances, in the context of an increasingly complicated world.

As a practicing clinical psychologist, I was not very familiar with the writings of Gilbert Gottlieb when I was invited to contribute to this book. Having now read some of his work (e.g., Gottlieb, 1998), and what others have written about it (Greenberg, 2007), I understand what prompted the invitation. Although I do not use the terminology Gottlieb employed (e.g., probabilistic epigenesis, bidirectionality), I am in agreement with him on the meaning of it and of its implications for
behavioral development. Like Gottlieb, I am critical of the central dogma of molecular biology. I believe that environmental factors play a crucial role in human development and behavior. On the other hand, genetic researchers frequently argue that genes play a predominant role. However, we have seen that the research supporting this position is greatly flawed by factors that Gottlieb also discussed, not the least of which is the bias of researchers and interest groups seeking to promote the primacy of genetics. Obviously, the development of every organism is influenced by heredity and environment, but this was understood long before behavioral genetic and psychiatric genetic research came onto the scene.

It is clear that theories of human psychological trait variation and psychopathology that emphasize genetics are harmful in that they serve to divert society’s attention from the need to change the environment. My work is not really about fusing nature and nurture, as it was for Gottlieb, but is mainly about illuminating the fatal flaws of behavioral genetic and psychiatric genetic research, regardless of the role that genes might play in the development of particular traits. The current emphasis on genetics helps absolve society from the responsibility of making the social and political changes necessary to improve the human condition. I trust that readers will see the relationship between my ideas and those Gottlieb put forward.

Regardless of possible genetic influences, it is clear that factors such as culture, family, birth cohort, political policies, access to health care, nutrition, religion, education, the mass media, and oppression on the basis of race, sex, sexual orientation, or social class, play a crucial and dominant role in shaping who we are and how we function and behave. Unfortunately, genetic research often diverts society’s attention from identifying and mitigating critical environmental factors that cause human distress and disease, and impede human growth (Chaufan, 2007). It contributes to putting off the day when society decides to undertake a serious effort to implement the necessary environmental interventions to alleviate and prevent human suffering, and to promote human growth to the fullest.

Acknowledgments

I would like to thank Galina Gerasimova, Claudia Chaufan, and Carl Ratner for providing helpful feedback on an earlier version of this chapter. The views put forward in this chapter do not necessarily reflect those of the reviewers, and any errors are the responsibility of the author.

Notes

1. Frank Miele (2005), a supporter of behavioral genetics, quoted Plomin et al.’s 1998 conclusions in his rejoinder to my (2005a) description of these researchers’ negative
results. Although Miele strongly implied that Plomin et al.'s conclusions trumped their data, he actually lent support to critics’ longstanding contention that genetically-oriented researchers' beliefs and biases strongly influence their research, and strongly influence the conclusions they draw from their research. Apparently, Plomin and his colleagues were committed to conclude in favor of genetics regardless of what their findings showed.

References


