Twenty-Two Invalidating Aspects of the Minnesota Study of Twins Reared Apart (MISTRA)

(FULL VERSION)

By Jay Joseph, Psy.D.

November, 2018

Author e-mail: jayjoseph22@gmail.com

Table of Contents

Tables and Figures

Summary

The 1979-2000 "Minnesota Study of Twins Reared Apart" (MISTRA) by Thomas J. Bouchard, Jr. and colleagues is often cited in support of the claim that genetic factors play a major role in causing differences in human behavior such as IQ and personality. Because the study used reared-apart monozygotic (identical) twin pairs ("MZA" pairs), who are genetically identical and supposedly experienced no environmental influences in common, it is widely seen as having provided definitive evidence in favor of the nature side of the "nature-nurture" debate. The present analysis closely examines how the MISTRA researchers arrived at their conclusions, and shows that these conclusions depended on the acceptance of many questionable or false assumptions, some of which Bouchard recognized "are likely not to hold." The researchers' strong biases in favor of genetic explanations had a major impact on the methods and comparisons they used, and on the conclusions they reached. Remarkably, they omitted their reared-apart dizygotic (fraternal) twin pair ("DZA" pair) correlations from their IQ study, even though they previously had designated DZA pairs as the MISTRA control group. Based on the *near* full-sample correlations that have been published, the MISTRA MZA and DZA group IQ correlations did not differ at a statistically significant level, which the study required as an important step in the process of determining whether genetic factors influence IQ scores. To this day, the MISTRA full-sample DZA group IQ correlations have not been published. At the same time, the researchers have prohibited independent analysts from inspecting and reviewing the MISTRA raw data. Twenty-two major problem areas in the study are discussed, including that the methods used to obtain the MZA sample led to the inclusion of behaviorally similar pairs, and that most pairs used in studies of this type were only *partially* reared apart. In addition, even perfectly separated MZA pairs experience many non-familial behavior-shaping environmental influences in common. The researchers, however, either denied that such environmental influences exist, or counted them as genetic influences. This created a genetic "heads I win, tails you lose" type of study that guaranteed that genetic interpretations of above-zero MZA group psychological test-score correlations would prevail. Highly publicized anecdotal stories of individual reunited twin pairs provide no evidence in support of genetic theories of human behavior. These selectively reported stories have been used mainly to sell such theories to the general public. Due to the many environmental confounds and other types of biases found in the study, most of which are also found in the "twins reared apart" studies that came before it, the MISTRA was unable to disentangle the potential influences of genes and environments. Therefore, its findings in favor of genetic influences on human behavioral differences, major or otherwise, must be rejected.

Abbreviations: **16PF** = 16 Personality Factor Questionnaire; **CAP** = Colorado Adoption Project; **CPI** = California Psychological Inventory; **DZA** = dizygotic (fraternal) twins reared apart; **DZT** = dizygotic (fraternal) twins reared together; **IQ** = intelligence quotient; **MISTRA** = Minnesota Study of Twins Reared Apart; **MMPI** = Minnesota Multiphasic Personality Inventory; **MPQ** = Multidimensional Personality Questionnaire; **MZA** = monozygotic (identical) twins reared apart; **MZT** = monozygotic (identical) twins reared together; **SES** = socioeconomic status; **TRA** = twins reared apart (study); **WAIS** = Wechsler Adult Intelligence Scale.

"Data don't tell stories, scientists tell stories"

—An academic supervisor, quoted in cognitive neuroscientist Chris Chambers' 2017 book *The Seven Deadly Sins of Psychology*[1]

"There is a danger of concealing assumptions which have no factual basis behind an impressive façade of flawless algebra"

—British scientist Lancelot Hogben, 1933[2]

"The usefulness and generalizability of the findings from [twins reared apart] studies depend, as they do for all research designs, on how well the assumptions of the design are met"

—Minnesota twin researcher Thomas J. Bouchard, Jr., 1998[3]

The Minnesota Study of Twins Reared Apart

Headlines and book titles send strong messages to the public about the reported findings from twin research: "Twins Separated at Birth Reveal Staggering Influence of Genetics," "How Genes Shape Personality," "Born That Way," "Are We Hardwired?," "Mean Genes," "DNA and Destiny," "Twin Brothers Separated at Birth Reveal Striking Genetic Similarities," "Study Raises the Estimate of Inherited Intelligence," "Life of Crime Is in the Genes, Study Says," "The Gene Bomb," and so on. Pulitzer Prize winning author Lawrence Wright, in his 1997 book about twin studies and behavioral genetic research *Twins: And What They Tell Us About Who We Are*, wrote that "the science of behavioral genetics, largely through twin studies, has made a persuasive case that much of our identity is stamped on us from conception; to the extent that our lives seem to be pre-chosen—all we have to do is live out the script that is written in our genes."[4] The potential impact of these messages on social relations and political policies is enormous.

I have dedicated much of the past two decades to critically examining behavioral genetic and psychiatric genetic theories, which hold that genetic factors play an important role in causing human behavioral differences.[5] Given the lack of gene discoveries, these theories continue to be based on the results of family, twin, and adoption studies. A sizable portion of my 2015 book *The Trouble with Twin Studies: A Reassessment of Twin Research in the Social and Behavioral Sciences* looked into the potentially invalidating problems of so-called "separated" or "twins reared apart" (TRA) studies, with a special focus on the famous "Minnesota Study of Twins Reared Apart" (MISTRA).[6] I concluded that, like studies of reared-together twins, the MISTRA and other TRA studies had failed to produce scientifically acceptable evidence that genetic factors play a role in causing differences in human cognitive ability (or "intelligence,"

supposedly measured by IQ tests), "personality," and human behavior in general. I summarized the main problems in TRA research in a 2014 online article. Estimating the "heritability of IQ" has been the main area of TRA study focus, with personality and other types of behavior playing a secondary role.

The study was headed by University of Minnesota psychologist Thomas J. Bouchard, Jr., with Bouchard's psychologist colleagues Nancy L. Segal, David T. Lykken, Matt McGue, and Auke Tellegen playing important roles. The researchers reported results in areas that included IQ, personality, special mental abilities, substance abuse and antisocial behavior, medical characteristics, cardiac characteristics, religiosity, dental parameters, workplace values, reproductive outcome, sensation seeking and control, "morningness-eveningness," headaches, dietary preferences, authoritarianism, social attitudes, and reading comprehension.

The volunteer-based MISTRA is by far the most cited and discussed of the six TRA studies that have appeared since 1937. The study was largely financed (about 60% or $1.42 million, roughly $3 million in 2018 U.S. dollars) by the Pioneer Fund, an organization created in the late 1930s to support eugenics and racial differences research.[7] According to psychologist-historian William H. Tucker, who published a 2002 book about the organization, the Pioneer Fund

> "was created by American supporters of the Third Reich, men who launched a campaign to ship black citizens back to Africa and wished to emulate the [anti-Semitic] Nuremberg Laws in the United States…The fund has continued to provide the resources for scientists maintaining that African Americans are intellectually inferior and for researchers whose work is not directly associated with race but whose conclusions about the importance of individual genetic differences are nevertheless likely to be helpful to the fund's larger purpose" (italics in original; links added).[8]

"If not for Pioneer," Bouchard said in 2009, "we would have folded long ago."[9] Of the total MISTRA funding of $2.33 million (roughly $4.5 million in 2018 U.S. dollars), an additional $185,000 was supplied by the Koch Charitable Foundation.[10]

Impact of the Study

The MISTRA has been referred to as "arguably, the most famous social science project in the last quarter of the twentieth century."[11] According to MISTRA researcher Nancy Segal, in her 2012 book about the study, *Born Together—Reared Apart: The Landmark Minnesota Twin Study*, the study "forever changed the way people think about the roots of human behavior."[12] The public's knowledge of the MISTRA is based largely on textbook descriptions and the works of authoritative experts, television reports, numerous popular books, and countless print and online articles that have appeared since 1979. A major theme of these reports and publications has been that the researchers discovered, often to their amazement, that genetic factors play a major role in most aspects of human behavior and abilities. Most people, including most academics, rely on textbooks and other secondary source accounts of behavioral genetic research. These accounts, however, are often inaccurate, at times appear to be based on other secondary sources, and usually endorse the original researchers' conclusions with little critical analysis.

Regardless of the original investigators' intentions and beliefs, twin studies have been cited by various commentators as supplying "scientific evidence" in defense of economic inequality and the social status quo, in support of racism and other forms of oppression, as an explanation for socially disapproved behavior (such as criminality), and in support of cutting back or eliminating needed social programs. The MISTRA findings have been cited in support of genetic (biological) determinism, which predates twin research and refers to the belief that genetic factors play a predominant role in causing differences in human behavior and mental abilities, and that environmental factors play only a minor role, if any. According to this view, which historically has been promoted by the economically and politically powerful in support of their interests, social problems, poverty, and psychiatric conditions are mainly the result of bad heredity.

MZA, MZT, and DZA Twin Pairs

TRA researchers calculate mean (average) psychological test-score correlations (for example, IQ, personality) in a group of MZ (monozygotic, identical) twin pairs, who supposedly were separated near birth and grew up apart in different homes. These twins share 100% of their segregating genes, and are known as "monozygotic twins reared apart," or "MZA" pairs.[13] Identical twins reared together in the same home are known as "monozygotic twins reared together," or "MZT" pairs. In the MISTRA, MZA and MZT group test-score correlations were compared in order to assess the role of what behavioral geneticists call "shared" environmental experiences. I will explain shortly how the researchers arrived at their conclusions in favor of genetic influences, and how they estimated heritability.

The MISTRA also collected a sizable sample of reared-apart DZ twins (dizygotic, fraternal), who are known as "dizygotic twins reared apart," or "DZA" pairs. These pairs also were supposedly separated near birth and grew up apart in different homes, but like ordinary siblings they share on average only 50% of their segregating genes in common. As I will discuss later, from the beginning Bouchard designated DZA pairs as the MISTRA control group. The study was conducted between 1979 and 2000, continuing to add newly recruited pairs as the study progressed. The final 2000 sample consisted of 81 MZA and 56 DZA pairs. Of the 56 DZA pairs, 38 were same-sex, and 18 were opposite-sex.[14] Many academic publications based on the MISTRA data have appeared since 2000.

Important Terms and Concepts

I use the term "behavioral resemblance" to include two main areas. The first, or *anecdotal* area, refers to the observed behavioral similarity of individual twin pairs as reported by researchers, journalists, and others. For example, the members of a reunited twin pair who wear similar clothes and enjoy fishing. The second, and more important "*scientific*" area, refers to how twins correlate on psychological tests such as IQ, personality, special mental abilities, and vocational interest tests, and how researchers interpret these correlations.

A *correlation coefficient* measures the extent to which twins' scores vary together, and ranges from -1.0 (strong negative relationship), through 0.0 (no relationship), to 1.0 (strong positive relationship). Correlations assess relationship, but do not address what *causes* the relationship. This must be determined by other methods. For instance, a hypothetical sample of

50 Argentinian MZA pairs would probably find a perfect positive correlation (1.0) for the ability to speak the Spanish language, which wouldn't mean that there is a gene or a genetic predisposition for speaking Spanish. Twin researchers use the *intraclass correlation* method (ICC), which is often used when assessing how much relatives, in this case twin pairs, resemble each other with regard to certain characteristics or traits. Although TRA researchers often highlight behavioral similarities found in the anecdotal area, their main conclusions are based on how they interpret above-zero MZA group psychological test-score correlations found in the scientific area.

Robert Plomin and his behavioral genetics colleagues have defined *heritability* as "the proportion of phenotypic variance that can be accounted for by genetic differences among individuals."[15] The key word here is *variance*, which is a statistical measure of how much scores or observations differ from their average value. Genetic researchers believe that heritability, a concept that has been the subject of much controversy, indicates the extent to which variation among people is explained (accounted for) by genetic influences. Plomin and colleagues wrote that assessing the causes of variation allows researchers to estimate "how much genetics contributes to a trait," and allows them to quantify the "relative importance" of genetic and environmental influences.[16] Heritability estimates range from 0% to 100% (0.0 to 1.0).

An *assumption* is something taken for granted or accepted as true without proof. The project or investigation then treats it, and researchers arrive at conclusions, as if it were true. Whether an assumption is true or false can completely change the results of a study.

TRA Studies and the Twin Method

By far the most common way that twins have been used in behavioral genetic and psychiatric genetic research has been through the use of "classical twin method" studies that compare the behavioral resemblance of MZTs, versus a group of reared-together same-sex dizygotic twin pairs (DZTs). Genetic findings in these studies are based on the assumption that MZTs and DZTs grow up experiencing roughly equal environments. Critics have shown, however, that MZTs in fact grow up experiencing much more similar environments and treatments, much closer emotional bonds, and much greater levels of "identity confusion" than experienced by DZTs, and that the twin method's all-important MZT-DZT "equal environment assumption" (EEA), therefore, is false.[17] This means that, because the twin method is unable to disentangle the potential influences of genes and environments, the greater behavioral resemblance of MZT versus DZT pairs can be completely explained by non-genetic factors.

Taking note of the controversy surrounding the twin method, TRA researchers and their supporters have claimed that studies of reared-*apart* twins provide a more definitive method of separating potential genetic and environmental (nature and nurture) influences on behavior, and that the results from these studies have validated the twin method. TRA studies are few in number (six) and are very difficult to perform because twins are rarely separated in early life, and because it is difficult to identify, recruit, reunite, and study such pairs. *In theory*, because twins are genetically identical but grew up apart, TRA researchers were able to cleanly separate the potential influences of genes and environments on IQ and personality. As we will see, the reality was something very different.

The Famous MISTRA *Science* Article

The most famous and most cited of the many MISTRA publications was a 1990 article published in *Science*, one of the world's leading scientific journals.[18] This article was co-authored by Bouchard, Lykken, McGue, Segal, and Tellegen. The editor of *Science*, Daniel Koshland, had invited Bouchard to submit an article about the study and its results to the journal. In Bouchard's view, "this article legitimated the study."[19] The "results from the *Science* paper," Segal wrote, "appeared in hundreds of newspapers, magazines, and broadcasts across the country and around the world."[20] The complete summary (abstract) of the study reads as follows:

> "Since 1979, a continuing study of monozygotic and dizygotic twins, separated in infancy and reared apart, has subjected more than 100 sets of reared-apart twins or triplets to a week of intensive psychological and physiological assessment. Like the prior, smaller studies of monozygotic twins reared apart, about 70% of the variance in IQ was found to be associated with genetic variation. On multiple measures of personality and temperament, occupational and leisure-time interests, and social attitudes, monozygotic twins reared apart are about as similar as are monozygotic twins reared together. These findings extend and support those from numerous other twin, family, and adoption studies. It is a plausible hypothesis that genetic differences affect psychological differences largely indirectly, by influencing the effective environment of the developing child. This evidence for the strong heritability of most psychological traits, sensibly construed, does not detract from the value or importance of parenting, education, and other propaedeutic interventions."[21]

I will refer to this summary (abstract) at various points, but right up front we see that Bouchard and colleagues did not say how they determined that "70% of the variance in IQ was found to be associated with genetic variation," nor did they clearly state what conclusions followed from the finding that the MZA and MZT behavioral correlations were similar. Bouchard and colleagues believed that they had found "evidence for the strong heritability of most psychological traits," but in the summary they did not say how they arrived at this conclusion. Since the finding that the MZA and MZT correlations were similar was the only twin-based result reported in the summary, one might conclude that this result led to the conclusions in favor of the "strong heritability of most psychological traits," and to the 70% IQ heritability estimate. As we will see, this was not the case.

Bouchard and colleagues noted that "the study of IQ is paradigmatic of human behavior genetic research," and that "IQ has been at the center of the nature-nurture debate."[22] The two main IQ measures they used were the "Wechsler Adult Intelligence Scale" (or "WAIS") and the Raven's Progressive Matrices /Mill-Hill Vocabulary Scale composite. A third way that cognitive ability was evaluated was the "First Principal Component of Special Mental Abilities," which Segal defined as "a variable subsuming a larger number of correlated variables" based on two special mental ability batteries. As Segal described it, "the WAIS is an individually administered test consisting of six verbal subtests (Information, Digit Span, Vocabulary, Arithmetic, Comprehension, and Similarities) and five nonverbal subtests (Picture Completion, Picture Arrangement, Block Design, Symbol, and Object Assembly)."[23] Many leading IQ researchers view the Raven test, which is non-verbal and involves problem-solving skills, as a highly "*g*-

loaded" test that is not biased by the culturally loaded questions found in standard IQ tests such as the WAIS and the Stanford-Binet. Some critics of the Raven test, and of IQ tests in general, argue against this claim and against the concept of "general intelligence."

The 1990 MISTRA sample consisted of 56 MZA and 30 DZA pairs. The *Science* article reported what Segal called the "long awaited IQ data."[24] The MZA group intraclass correlations for the WAIS, Raven/Mill-Hill, and First Principal Component IQ measures were .69, .78, and .78 respectively. MZA correlations for the two reported "special mental abilities" tests (verbal, perceptual, spatial, and memory tasks) were .45 and .48, and for the two reported personality inventories (tests) they were .48 and .50. They also reported correlations for "psychological interests" and "social attitudes." These results are seen in Table 1. As I will discuss later, Bouchard, Segal and colleagues did not provide any DZA group results or correlations in this 1990 *Science* article, even though they had designated DZAs as the MISTRA control group.

Table 1

Psychological Test-Score Correlations for Selected Behaviors: The 1990 MISTRA *Science* Article

Test	MZA *r*	MZT *r*	DZA Control Group *r*
WAIS Full Scale IQ	**.69**	**.88**	**Not reported**
Raven, Mill-Hill Composite IQ	**.78**	**.76**	**Not reported**
First Principle Component IQ	**.78**	**"Not available"**	**Not reported**
Hawaii Special Mental Abilities	**.45**	**"Not available"**	**Not reported**
Comprehensive Special Mental Abilities	**.48**	**"Not available"**	**Not reported**
Multidimensional Personality Questionnaire	**.50**	**.49**	**Not reported**
California Psychological Inventory (personality)	**.48**	**.49**	**Not reported**
Strong Campbell Interest Inventory	**.39**	**.48**	**Not reported**
Religiosity Scales	**.49**	**.51**	**Not reported**
MPQ Traditionalism Scale	**.53**	**.50**	**Not reported**

Data from Bouchard, Lykken, McGue, Segal, & Tellegen, 1990, *Science, 250,* Table 4, p. 226. *r* = intraclass correlation; MZA = monozygotic twins reared apart; MZT = monozygotic twins reared together; DZA = dizygotic twins reared apart; MISTRA = Minnesota Study of Twins Reared Apart; WAIS = Wechsler Adult Intelligence Scale; MPQ = Multidimensional Personality Questionnaire. The 1990 MISTRA sample consisted of 56 MZA and 30 DZA pairs. The listed MZA correlations were based on fewer reported pairs, ranging from 31 to 52 MZA pairs. No DZA correlations of any kind were reported in the 1990 MISTRA *Science* article. "Not available" status was reported by the researchers.

Because the researchers based their *Science* study conclusions on the claim that the MZA correlations seen in Table 1 "directly estimate heritability" (see below), MISTRA heritability estimates were similar to the MZA correlations for the behavior in question. This placed the 1990 *Science* article estimated IQ heritability at about 70%, while personality heritability was about 50%. In a 1993 publication the researchers set IQ heritability at 75%, and in another 1993 publication Bouchard placed the personality heritability estimate in the 40-50% range.[25] Bouchard and colleagues concluded in their *Science* article that "general intelligence or IQ is strongly affected by genetic factors," and that in general, genetic factors "exert a pronounced and pervasive influence on behavioral variability."[26] In a later publication, Bouchard wrote that the MISTRA results helped show that the finding of a "large degree of genetic influence" on differences in IQ scores is "irrefutable."[27]

In the 1990 *Science* article, in Segal's *Born Together—Reared Apart*, and in other MISTRA publications, the MISTRA MZA group IQ correlations were presented as being "remarkably consistent" with the findings from the three original TRA studies. These studies were published in 1937, 1962, and 1965, respectively by Horatio Newman and colleagues in the United States (19 MZA pairs), James Shields in England (44 MZA pairs), and Niels Juel-Nielsen in Denmark (12 MZA pairs).[28] (The IQ TRA study published by British psychologist Cyril Burt had been discredited.) Shields was the only original researcher to study DZA pairs, although his original plan had been to exclude such pairs. His sample of 11 pairs did not figure into his major conclusions.

The MISTRA researchers discussed the findings from these earlier TRA studies as if they were valid studies of reared-apart twins, when in fact they were riddled with numerous invalidating errors and biases. Continuing the work of earlier critics, I discussed these invalidating errors and biases in the first three chapters of *The Trouble with Twin Studies*. The MZA pairs found in these studies were, for the most part, only *partially* reared apart.

Here I will show that the *Science* article and other MISTRA publications were subject to numerous problems and biases that are rarely mentioned in mainstream accounts. I will describe the main problem areas, expanding on and adding to the points I raised in *The Trouble with Twin Studies*. I listed many of these problem areas very briefly in a 2018 tribute to the pioneering critic of TRA studies, the psychologist Leon J. Kamin (1927-2017).

The huge impact of the MISTRA, in addition to the harmful and regressive genetic determinist social and political policy implications that flow from its conclusions, necessitates a detailed analysis of the science behind the study's major claims and conclusions as found in the 1990 *Science* article and in other original peer-reviewed MISTRA publications, and in Segal's *Born Together—Reared Apart*.

MISTRA "Model-Fitting" Analyses

Since the 1970s, behavioral genetics has embraced "biometrical model fitting" statistical analyses. Model fitting, according to the leaders of the field, is a "technique for testing the fit between a model of genetic and environmental relatedness against the observed data. Different models can be compared, and the best-fitting model is used to estimate genetic and

environmental parameters."[29] Model-fitting analyses attempt to partition (A) genetic, (C) "shared environment," and (E) "non-shared environment" contributions to behavioral variation in a population. Heritability estimates are based on the genetic "A" contribution.

Segal described the MISTRA basic model, which was based on the assumptions that "shared genes underlie similarity between relatives, mating occurs at random (is not assortative), genetic effects are additive, genetic and environmental effects are independent from each other, and genetic and environmental effects combine additively."[30] The model's assumption that "shared genes underlie similarity between relatives" is another way of saying that the model assumes that shared environmental factors *do not* underlie similarity between relatives.

Contrary to model-fitting assumptions, critics have argued convincingly that genetic and environmental effects *are not* independent from each other, and that gene-environment interactions reduce or even invalidate heritability estimates produced by model-fitting analyses. Furthermore, we will soon see that there are many non-familial environmental influences and cohort effects that contribute to above-zero MZA behavioral correlations. The MISTRA model, however, is based on the assumption that no such influences exist.

DZA Pairs as the MISTRA Designated Control Group

Bouchard designated DZA pairs as the MISTRA control group at the beginning of the study, presumably to be compared with the MZA experimental group. According to Segal, "Bouchard's decision to use DZA twins as controls was made in a very early memo [dated March 5th, 1979[31]] to the 'Twin Research Team.' This was an important methodological improvement over past projects."[32]

In a 1986 publication, Bouchard, Segal, and colleagues wrote, "Our study is the first to have included a control group of dizygotic twins reared apart (DZA)." They emphasized that by using a DZA control group, they would be able to test genetic versus non-genetic explanations of above-zero MZA group psychological test-score correlations:

> "DZA twins allow us to test the two most common competing hypotheses proposed as alternatives to the genetic hypothesis as an explanation of the similarity between MZA twins: placement bias and recruitment bias."[33]

However, the MISTRA model-fitting procedure, which the researchers used for the first time two years later in their first major publication, a 1988 personality study by Tellegen, Bouchard, Segal and colleagues, was based on the assumption that "all resemblance between reared apart relatives is because of shared genetic factors."[34] The study was now based on the built-in model-fitting assumption that placement and recruitment biases *did not* influence MZA behavioral resemblance. This meant that instead of following their 1986 plan to use DZA pairs to *test* the above-mentioned "competing hypotheses," by 1988 the researchers simply *assumed* that the "genetic hypothesis" of the MZA group behavioral resemblance was correct, and they assumed that placement, recruitment, and all other environmental hypotheses were *in*correct. As Tellegen, Bouchard, Segal and colleagues wrote in this 1988 MISTRA personality study:

> "Although intraclass correlations are often informative and are presented here, for analytic purposes they can be misleading if MZ and DZ variances differ. Biometric [model-fitting] geneticists…therefore prefer analyzing variances over correlations. We also take this approach in this article."[35]

Although Bouchard wrote in 1984 that "a model-fitting approach to family data…is far superior to the simple calculation of heritabilities," I am unaware of any MISTRA publication or document, appearing prior to the 1988 Tellegen et al. publication, in which the researchers stated that they intended to base their conclusions on model-fitting results.[36]

How the Researchers Concluded That Genetic Factors Play an Important Role

It is very important to understand and to examine closely the methods, steps, and the stated and unstated assumptions that social and behavioral science investigators use to arrive at their conclusions.

It is often mistakenly reported in the media, in textbooks, in popular works, and even at the University of Minnesota's "Minnesota Center for Twin and Family Research" website that the MISTRA genetic findings were based on MZA versus MZT group comparisons.[37] (I made the same mistake myself in some earlier publications.)[38] Instead, as described throughout Segal's *Born Together—Reared Apart*, the MISTRA team compared MZA and MZT correlations in order to assess the influence of *environmental* factors, using the questionable behavioral genetic distinction between "shared" and "non-shared" environments. As Segal wrote in a 2017 book on twins, "Comparing MZA and MZT twin pairs tells how much sharing an environment affects behavioral and physical development."[39] In their 1990 *Science* article, Bouchard and colleagues found that "adult MZ twins are about equally similar on most physiological and psychological traits regardless of rearing status [MZA or MZT]," but concluded from this only that "common rearing enhances familial resemblance during adulthood only slightly and on relatively few behavioral dimensions."[40] Even if this conclusion were true—and common sense assures us that it is not true—environmental factors unrelated to common rearing could still account for MZA behavioral resemblance. (Later I will discuss many of these *non-familial* behavior-shaping influences.)

This leads us to the crucial question of how the MISTRA researchers *did* reach their conclusions in favor of "pronounced and pervasive" genetic influences on IQ and other types of human behavior. In the four steps described below, and in the accompanying Figure 1, I discuss and show how they arrived at these conclusions. Because they did not explain this process clearly and consistently in their publications (which contributed to the mistaken descriptions mentioned above), I have constructed these steps on the basis of information extracted from the original MISTRA publications, in addition to basic principles of twin research, psychological testing, and statistics. Because in different publications the researchers chose different methods of arriving at heritability estimates at Step 3, these two "either/or" methods are described/shown as Step 3A and Step 3B.

Step 1 is usually achieved in TRA studies, and therefore has not been a major area of dispute between twin researchers and their critics. The main issue that *has* been disputed is the question of what factor or factors *cause* above-zero MZA group psychological test-score

correlations: genetic factors, non-genetic factors, or a combination of both. In addressing this key question, I will show that there are major problems with Steps 2, 3A, and 3B, which led the researchers to arrive at the mistaken Step 4 conclusion that genetic factors are an important cause of human behavioral variation.

The four main steps that the MISTRA researchers used to arrive at their conclusions in favor of genetics are described below. As we will see, MZT correlations were not part of the process that led them to these conclusions (other than possibly being included in model-fitting analyses). In the accompanying Figure 1 and in the discussion that follows, we will see that the researchers bypassed Step 2 and Step 3A when assessing their IQ (general cognitive ability) results. The quotations found in some of the step descriptions are taken from leading MISTRA publications.

Step 1: MZA correlation > zero

The mean (average) MZA group psychological test-score correlation (e.g., IQ, personality) must be higher than zero (0.0) at a level that falls below the conventional .05 level of statistical significance. If not, the study finds no genetic influence on the behavior in question because, statistically speaking, the MZA group correlation is zero (there is no relationship between the twins' test scores).

Step 2: MZA > DZA

Because MZAs are more genetically alike than are DZAs (100% vs. an average 50%), the significantly above-zero MZA group correlation found in Step 1 must also be higher than the corresponding DZA control group correlation at a statistically significant level. If not, the study finds no genetic influence on the behavioral characteristic in question, which suggests that *non*-genetic influences alone were responsible for raising both the MZA and the DZA group correlations above zero. "The simple comparison of the MZ (or MZA) and DZ (or DZA) intraclass correlations is an important first step in behavioral-genetic analysis because this demonstrates *whether or not* there is genetic influence on the trait" (italics added).[41]

Step 3A: Model-fitting heritability (most non-IQ MISTRA studies)

Based on the assumption that "all resemblance between reared apart relatives is because of genetic factors," in addition to many other assumptions, statistical "model-fitting" analyses use MZA, DZA and other data to produce sizable heritability estimates. If environmental effects are found to influence MZA group psychological test-score correlations, these are counted as *genetic* effects based on the claim that twins create more similar environments for themselves because they behave more similarly for

genetic reasons: "The immediate causes of most psychological variations are probably environmental in nature. However, the environments of individuals are significantly fashioned by their genotypes that selectively guide them toward certain people, places, and experiences and away from others."[42] The researchers assume that the heritability estimate (ranging from 0% to 100%) indicates the proportion of the population variance explained by genetic factors for the behavioral characteristic in question, and they assume that such estimates also indicate the degree to which the characteristic is influenced by genetic factors.

OR

Step 3B: The MZA correlation directly estimates heritability (IQ, 1990 *Science* article)

The MZA group correlation is assumed to directly estimate heritability, because it is assumed that MZA pairs share only their genes in common: "The MZA intraclass correlation directly estimates broad heritability because MZA twins share all their genes but do not share their rearing environment. In other words, MZA co-twins have only their genes in common, so their observed similarities reflect their shared genes."[43] Like model fitting procedures, environmental influences on MZA correlations are counted as genetic influences. Because heritability is estimated directly and solely from the MZA group correlation, DZA control group correlations are ignored, omitted, or both.

Step 4: Conclusion

After finding "evidence for the strong heritability of most psychological traits," it is concluded that genetic factors "exert a pronounced and pervasive influence on behavioral variability."[44]

I will refer to these steps throughout this article. The researchers' decision-making process is shown in Figure 1.

Figure 1

The MISTRA Data Interpretation Process: Genetic Findings

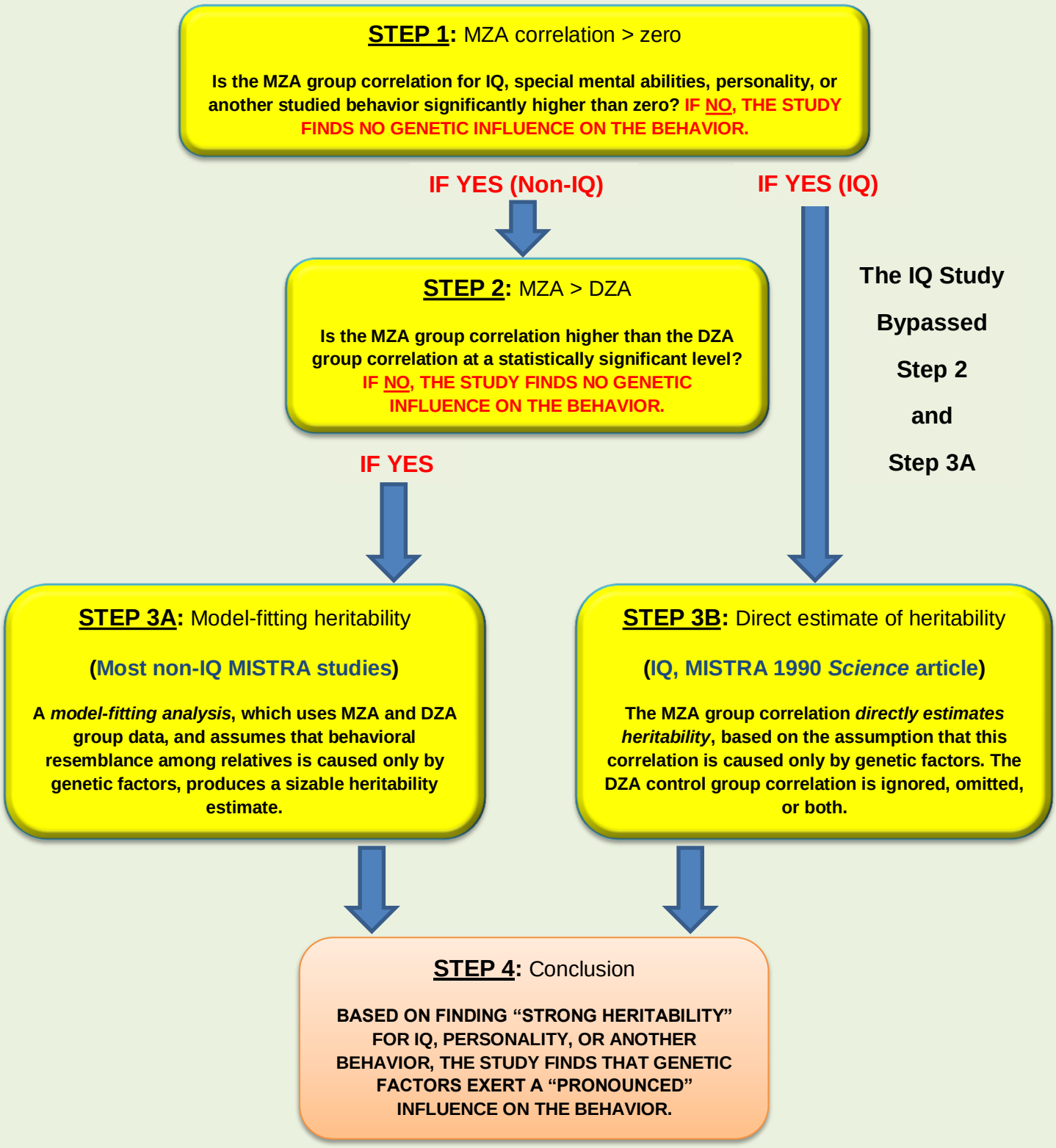


Figure 1 provides a roadmap of the decision-making process that the researchers themselves did not provide. The logic that they used to arrive at their conclusions in favor of genetics was inconsistent and difficult to follow. The summary (abstract) of the 1990 *Science* article could be interpreted as saying that the study's 70% IQ heritability estimate was based on MZA-MZT comparisons, but in the body of the article this estimate was based only on the "MZA correlation directly estimates heritability" claim (Step 3B). Turning to non-IQ behaviors such as vocational interests and personality, in most MISTRA publications the genetic findings were based on Step 3A model-fitting results, yet in the 1990 *Science* article, genetic findings were based only on Step 3B. Furthermore, the *Science* IQ heritability estimate was arrived at "under the assumption of no [MZA] environmental similarity," yet later in the article the researchers concluded that "MZA twins are so similar in psychological traits because their identical genomes make it probable that their effective environments are similar." Apparently, MZA twins *did* experience environmental similarity.

The 1990 *Science* article began with the statement, "Monozygotic and dizygotic twins who were separated early in life and reared apart (MZA and DZA twin pairs) are a fascinating experiment of nature. They also provide the simplest and most powerful method for disentangling the influence of environmental and genetic factors on human characteristics." However, this "fascinating" and "most powerful" "experiment of nature" did not appear in the article, and all DZA findings and correlations were omitted! The summary stated that the study tested "more than 100 sets of reared-apart twins or triplets," but the article reported correlations for less than half of these sets. It is simply amazing that this confusing (and confused) article based on the obviously false MZA "no environmental similarity" assumption (see below), where the DZA *control group* correlations were omitted, made it through the supposedly rigorous peer-review process of one of the world's leading scientific journals, and continues to be cited favorably in leading psychology textbooks and elsewhere as a "landmark study."

The Disappearing DZA Group IQ Correlations

We have seen that, depending on which behavioral characteristic they were studying, the researchers used different methods to arrive at their major conclusions. Apart from the 1990 *Science* article, for most *non-IQ* MISTRA-studied behaviors they appear to have followed Steps 1, 2, 3A, and 4, assuming that the Step 2 MZA-DZA group correlation comparisons were incorporated into the model-fitting analyses.

We have also seen that when analyzing their *IQ data* in the 1990 *Science* article and elsewhere, the researchers did not compare MZA and DZA correlations, and they did not use model-fitting analyses. Instead, as seen in Figure 1, they bypassed Step 2 and Step 3A, and based their conclusions on the Step 3B assumption that the MZA group correlation directly estimates heritability because MZAs share only their genes in common, and that environmental influences count as genetic influences. By using this maneuver, Bouchard and colleagues ***completely removed the DZA control group data from the process they used to arrive at their conclusions in favor of "pronounced" genetic influences on IQ.***

Although in 1997 Bouchard recognized that "formal testing of substantive hypotheses via model fitting has now become the norm in behavior-genetic research," he and his colleagues never published a model-fitting analysis of the MISTRA IQ data, nor did they publish their full-sample DZA group IQ correlations.[45] We will see that this may have occurred because, as the *near* full-sample DZA IQ correlations published in 2007 and 2012 strongly suggest, the full-sample MISTRA MZA group IQ correlations were not higher than the corresponding DZA correlations at a statistically significant level, and therefore failed to meet the Step 2 requirement of finding a significantly higher MZA versus DZA correlation. As seen in Figure 1, the researchers bypassed two key steps to get around this problem, which enabled them to reach the standard behavioral-genetic conclusion that genetic factors have an important influence IQ scores.

"Hidden Flexibility" in Behavioral Research

Under past and current systems, researchers have "degrees of freedom" that give them the "hidden flexibility" to change various aspects of their study after reviewing the data, but before submitting their paper for peer review and publication.[46] In his 2017 book *The Seven Deadly Sins of Psychology: A Manifesto for Reforming the Culture of Scientific Practice*, cognitive neuroscientist Chris Chambers wrote that a "key feature" of researchers' decisions "is that they are hidden and never published."[47] Surveys suggest that "questionable research practices" are common in American psychological research.

Chambers showed that "hidden" manipulation occurs frequently by researchers attempting to obtain desired results, or to produce findings that will be accepted for publication by leading journals. Although a "cardinal rule in experimental design" is "that any decision regarding the treatment of data must be made prior to an inspection of the data," it is difficult for outsiders to determine whether this occurred, because social and behavioral science research does not yet require accountability through research pre-registration.[48]

Because, with the data in hand, the MISTRA researchers used their "hidden flexibility" to bypass Steps 2 and 3A in their IQ study—and to thereby ignore their DZA control group IQ correlations—their conclusions in this area rested on the assumption that all factors influencing significantly above-zero MZA group correlations should be counted as genetic factors, and on the accompanying assumption that the MZA correlation "directly estimates heritability" (Step 3B). As we will soon see, and as I argued in *The Trouble with Twin Studies*, the numerous biases built into the study, in addition to the many non-familial environmental influences experienced by MZA pairs, show that these assumptions are utterly false.

***

The all-important question to be answered in a TRA study is the following one:

Assuming that the tests and the behavioral categories in question are valid, what is the cause, or what are the causes, of statistically significant above-zero MZA group psychological test-score correlations?

In this article I present the critics' case against genetic interpretations of the MISTRA MZA group correlations in the form of 22 reasons (either standing alone or in combination with other reasons) why such interpretations should be rejected. Themes I will address include (1) bias in the twin sample and in the methods used to assess twins; (2) strong researcher bias in favor of genetic explanations of the results; (3) conclusions that were based on questionable or false assumptions and concepts; (4) a lack of accountability and transparency; and (5) that environmental (non-genetic) factors plus research bias plausibly explain the MISTRA results, especially in light of the failure to make confirmed discoveries of genetic variants that cause differences in IQ, personality, and other forms of behavior.

Similar to other behavioral genetic research methods and previous TRA studies, I will show that the MISTRA was unable to disentangle the potential influences of genes and environments (nature and nurture) on human behavior.

Twenty-Two Reasons to Reject the MISTRA Researchers' Conclusions in Favor of "Pronounced and Pervasive" Genetic Influences on Human Behavioral Differences

1. **TRA Studies Based on Volunteer Twins Recruited Through Media Appeals Produce Samples That Favor the Inclusion of Behaviorally Similar Twin Pairs.** We have seen that the original three TRA studies were published by Newman and colleagues (19 MZA pairs), Shields (44 MZA pairs), and Juel-Nielsen (12 MZA pairs). The MISTRA was subject to several problem areas that Kamin had identified in the Newman and Shields studies, which were based on volunteer twin pairs recruited through media appeals.[49]

   Kamin showed that in studies that recruited twins through the use of such appeals, the twins had to have been aware of each other's existence to be able to respond. They may have responded to the appeal, or they may have discovered each other, *because of* their behavioral similarities. Volunteer-based studies such as the MISTRA, as other leading behavioral geneticists have noted, "typically relied on identification by third parties or response to media appeals. Pairs may have come to the investigator's (and to each other's) attention because of their remarkable similarity."[50] Newman and colleagues recognized that in their volunteer-based 1937 study, "It seems possible that our [MZA] group is *more heavily weighted with extremely similar pairs* than with identical twins of less striking similarity" (italics added), which they viewed as an unintended bias resulting from the methods they used to recruit separated twin pairs.[51]

   A much less biased method of obtaining MZA pairs is to identify them through the use of population registers, yet only in Juel-Nielsen's 1965 study were some MZA pairs (8/12) identified in this way. According to Segal, "Ideally, researchers would want to find separated twin sets in national population registries that record all occurrences of multiple births and rearing circumstances. Such resources are available in the Scandinavian

countries….Lacking a national registry, the MISTRA studied a collection of cases because our reared-apart twins surfaced in many ways and at any time….The twins themselves and the people who heard about them contacted us because they knew about the MISTRA largely through the media attention the study had attracted."[52]

Shields saw the use of volunteers as a "risky procedure in most types of research, but inevitable" in his study.[53] And Lykken recognized that it "is well established that volunteers tend to be more intelligent than nonvolunteers."[54] Lykken also recognized that volunteer subjects are better educated and frequently hold middle-class values, which will further increase twins' behavioral resemblance for non-genetic reasons. Additional studies have shown that research volunteers share several other psychological and behavioral characteristics.[55] Samples based on volunteer twins also tend to be overrepresented by female pairs, which is reflected by the fact that the final MISTRA MZA sample was 60% female, while the same-sex DZA sample was 68% female.[56]

In her exhaustive 1981 book *Identical Twins Reared Apart: A Reanalysis*, psychologist Susan Farber observed that "approximately 90 percent of the known cases of separated MZ twins have been studied precisely because they were so alike," and that conclusions about their similarity were based on "circular reasoning."[57] Farber estimated that there were about 600 adult MZA pairs then living in the United States. Roughly one-half of the 81 MISTRA MZA pairs lived in the U.S., making it likely that, due to recruitment bias, the MISTRA U.S. sample was drawn from the most behaviorally similar 10% of the MZA population.[58]

The MISTRA researchers attempted to minimize ascertainment bias by "vigorously" recruiting all pairs they became aware of, regardless of whether they were MZA or DZA. They believed that this method created a sample that was not subject to such bias.[59] While this practice certainly was an improvement over the earlier studies (particularly the 1937 Newman study), the MISTRA sample remained substantially biased in favor of behavioral similarity.

In summary, the MISTRA's use of volunteer twin pairs biased the study in favor of the recruitment of MZA pairs who were more behaviorally similar when compared with randomly selected pairs drawn from the entire population of MZA pairs.

2. **"Heritability" Is One of the "Most Misleading [Terms] In the History of Science."** A major aspect of behavioral genetic studies, including the MISTRA, is the production of heritability estimates. The heritability concept was developed in the mid-20th century to help predict the results of selective breeding programs of farm animals, but was extended by behavioral geneticists and others into a measure of the degree of genetic influence on psychiatric disorders, and behavioral characteristics such as IQ and personality.[60] Heritability estimates are derived from correlations among relatives, or are produced by model-fitting analyses that often incorporate twin data.

The validity of the heritability concept and accompanying heritability estimates, however, has been disputed for decades. Critics have argued convincingly that heritability estimates do not and cannot indicate the "degree of genetic influence" on psychiatric disorders and

behavioral characteristics, nor do they indicate to what degree a behavioral characteristic is changeable.[61] Heritability is an attempt to explain variation, and does not indicate the "strength" or "weakness" of the presumed genetic influence—or by implication the strength or weakness of the environmental influence.

According to psychologist David Moore and his co-author David Shenk, "The term 'heritability'…is one of the most misleading in the history of science," because "contrary to popular belief, the measurable heritability of a trait *does not* tell us how 'genetically inheritable' that trait is. Further, it does not inform us about what causes a trait, the relative influence of genes in the development of a trait, or the relative influence of the environment in the development of a trait" (italics in original).[62] Some have argued that a heritability estimate serves no valid purpose apart from its original use as a predictor of the results of a selective breeding program.[63] It follows that MISTRA heritability estimates have little meaning, and do not indicate the "strength" of the claimed genetic contribution to the behavioral characteristic under study. By 2009, Bouchard and other leading behavioral geneticists recognized that "specific estimates of heritability are not very important."[64]

3. **The Researchers Refused to Share Their Raw Data and Information with Potential Critics.** When Kamin contacted Bouchard and requested access to the MISTRA raw data, Bouchard denied him access even under conditions where twin pairs would be identified only by code numbers, and where information about age would be omitted to guarantee non-identification.[65] As Bouchard told a journalist, he "wouldn't let Leon Kamin anywhere near" the MISTRA raw research material, although he was available to answer a "legitimate question."[66] Bouchard also turned down the separate requests by William Tucker and neuroscientist Steven Rose to review the raw data.[67] This "data hoarding" strategy, which violates the "Ethical Principles" of the American Psychological Association, violated basic scientific principles, especially since TRA studies are nearly impossible to reproduce (replicate) due to the increasing rarity of separated twins.[68] Bouchard denied Kamin and others the opportunity to inspect the raw data, and to possibly arrive at a different set of conclusions than arrived at by the genetically oriented MISTRA researchers.

A committee established to investigate the fraudulent activities of Dutch psychologist Diederik Stapel, and to make recommendations to prevent fraudulent research in psychology in the future, wrote in 2012 that "it must always remain possible for the conclusions to be traced back to the original data." To reduce the possibility of misconduct in research, they recommended (1) that data must be made available to other scientists, (2) that data should be stored in a way that ensures that it cannot be modified, (3) that someone should be designated as being responsible for safeguarding the data, and (4) that publications should be required to state where the data is located and how it can be accessed. The committee concluded that academic "journals should only accept articles if the data concerned has been made accessible in this way."[69]

The MISTRA researchers, on the other hand, denied access to Kamin and other potential critics, allowing access only to friendly colleagues who supported their work and

conclusions.[70] In *Born Together—Reared Apart*, Segal wrote that the "burden of proof lies with the critics" to show that the study contained invalidating biases. "Bias must be demonstrated, not assumed," she wrote.[71] This placed critics in a classic "Catch-22"-like position because, if they were known to be inclined to look for bias, the Minnesota researchers denied them access to the raw data.[72]

When researchers fail to make their raw data available for inspection and analysis by qualified reviewers, we must automatically reject their findings when they are based on samples taken from rare populations that cannot be reproduced. Amazingly, the MISTRA researchers ask us to accept their conclusions, most of which have important social and political policy implications, based on data that they will not allow anyone to see, other than people who are likely to agree with them. The study's findings should be rejected for this reason alone.

4. **Conclusions Were Based on Assuming the Validity of What Are in Fact Disputed Psychometric Tests and Concepts.** The MISTRA findings were based on accepting the validity of concepts such as "IQ," "general intelligence (*g*)," "personality," "model fitting," and "heritability." These concepts, however, were and remain controversial, and the validity of the "heritability of IQ," the "heritability of personality traits," and the "heritability of psychiatric disorders" concepts are also controversial. If these concepts are invalid, the MISTRA and most other behavioral genetic studies are invalid as well.

5. **The Researchers' Conclusions Were Influenced by Their Strong Biases in Favor of Genetic Explanations.** Similar to many research projects in the social and behavioral sciences, the MISTRA researchers' conclusions were influenced by *confirmation bias*, which is the tendency for people (including researchers and their critics) to search for, interpret, favor, and recall information in ways that confirm their pre-existing beliefs or theories. In this case their methods and conclusions were heavily influenced by hereditarian biases, which favored genetic interpretations of the data.

The researchers operated under the assumption that genetic factors are an important cause of differences in IQ and behavior in general. If a correlation they calculated suggested otherwise, they assumed that something was wrong the correlation, and/or that the correlation could be explained on genetic grounds.

Bouchard had been an early supporter and teacher of Arthur Jensen's hereditarian IQ theories,[73] which drew protests from University of Minnesota students in the early 1970s.[74] In 1976, three years prior to initiating the MISTRA, Bouchard had written that "human intelligence," as supposedly measure by IQ tests, "is largely under genetic control," that social class differences in intelligence "have an appreciable genetic component," and that due to reproduction patterns, the possibility of a decline in national intelligence "should be subject to continual scrutiny."[75]

Bouchard wrote a positive 1995 review of Richard Herrnstein and Charles Murray's controversial and widely publicized 1994 book *The Bell Curve: Intelligence and Class Structure in American Life*, and endorsed its authors' "taboo" conclusion that genetic factors play a role in causing racial differences in IQ scores.[76] Bouchard also wrote an endorsement for the 2000 abridged edition of former Pioneer Fund Director J. Philippe Rushton's *Race, Evolution, and Behavior: A Life History Perspective*, where Rushton (1943-2012) argued that racial groups differ in inherited levels of intelligence, with people of African descent occupying the bottom position.[77] Using Pioneer Fund money, this book was mailed out unsolicited to tens of thousands of social scientists.[78]

MISTRA researcher David Lykken (1928-2006) proposed the establishment of "parental licensure" laws to help "cure" or reduce "black crime,"[79] which could include the forced "implantation of a long-acting antifertility drug."[80] He also wrote a very positive 2004 review of University of Ulster psychologist Richard Lynn's 2001 pro-eugenics book *Eugenics: A Reassessment*, which Lykken described as "an excellent, scholarly book, and one cannot reasonably disagree with him on any point unless one can find an argument that he has not already refuted."[81] In his 1995 book *The Antisocial Personalities*, Lykken endorsed the racial differences in IQ position of *The Bell Curve*,[82] and wrote that "one minor but significant casualty of the Hitler period was the loss to our language of the innocent and useful word, *eugenics*" (italics in original).[83]

The MISTRA researchers recognized no conclusions or interpretations other than their own, and Bouchard frequently portrayed scholars attempting to challenge his conclusions as the purveyors of "pseudoanalyses." Science historian Michael Rossi noted the earlier TRA researchers' reluctance to reach definitive conclusions in their studies, and contrasted this with the Minnesota researchers: "Didn't the Mistra scientists ever experience divergent interpretations of their own data? Didn't they ever argue among themselves over what their data meant? If not, on what basis did they conclude that their results made sense?"[84]

There is no indication that leading MISTRA researchers had any major disagreements or greatly differing perspectives. In a 2004 interview, Bouchard recalled that he had co-authored many papers with MISTRA colleagues David Lykken and Auke Tellegen, "and we never had a disagreement in 20 years."[85]

Despite Segal's belief that "science rests on data, not dialogue," results and data do not speak for themselves, and can be interpreted in many different ways.[86] Although like most behavioral genetic researchers Bouchard, Lykken, Segal, McGue and colleagues interpreted their results on the basis of their strong pre-existing biases in favor of genetics, this is not how most journalists and popular writers have presented the MISTRA story. MISTRA-friendly Pulitzer Prize winning cancer physician Siddhartha Mukherjee, to cite one example, wrote in his 2016 best-selling book *The Gene: An Intimate History* that "Bouchard's staff was repeatedly struck by the similarities between the twins."[87] The word "struck" suggests to unknowing readers that the researchers had no genetic biases as they observed and evaluated twins, and that they might even have had environmental biases. Rhetorical maneuvers of this type since 1979 have helped the MISTRA researchers' claims achieve greater legitimacy. As I have shown, *The Gene: An Intimate History* contained

many false or unsubstantiated statements about the study, and also presented schizophrenia twin research in a similarly inaccurate way.[88]

6. **The Researchers Published Only Minimal Information on the Twins.** Unlike the authors of the original three TRA studies, who provided a wealth of information about the twins they studied, the MISTRA researchers provided very little published case history information, and even failed to produce a table with basic demographic, degree-of-separation, and test-score information for each studied twin pair. Nancy Segal's 2012 book about the study provided an excellent opportunity to finally provide this information. However, she chose not to provide it and mainly reviewed previously published material, even though she had been given access to all files and documents associated with the MISTRA.[89]

   Following the publication of the MISTRA 1990 *Science* article, Harvard geneticist Jonathan Beckwith and his colleagues wrote a letter to *Science* criticizing the MISTRA team for "not publishing in a format that permits independent scrutiny. Investigators in this field should indicate the precise nature of being 'reared apart,' including, for example, whether it involved only being raised in separate households within the same community." Beckwith and colleagues wrote that "it is imperative that case studies be fully published."[90] In response, Bouchard and colleagues argued that it is "highly unlikely" that "the similarities between the MZA twins might be explained by unreported environmental similarities."[91] As evidence, they cited studies that found no IQ correlations between biologically unrelated individuals reared together, implying that common environment does not lead to IQ resemblance. Beckwith and colleagues appropriately requested much more information about the twins' life circumstances and degree of separation. Bouchard and colleagues responded to this request with (non-twin) correlation coefficients.[92]

   In a 1981 interview with Juel-Nielsen, Bouchard said that despite the apparent "neatness" of the TRA design, "when you end up working with these twins, and you start studying their lives in great detail, you really see how complicated the situation is. The apparent simplicity of the design is just overwhelmed by the complexity of individual lives."[93] This is undoubtedly true, but only Bouchard and his colleagues had access to the information showing the complexity of the twins' lives, and the complexity of the study itself. Perhaps independent reviewers, had they been allowed access to the raw data and information about the twins' life histories, would have had a different interpretation of the causes of the behaviors that resulted from the twin's complex lives.

7. **Most MZA Twins Were Abandoned Children, and the Generalization of TRA Findings is Questionable.** We must always remember that most MZA and DZA twins were adoptees who, as children, were abandoned by, or were taken from, their birthparent(s) for various reasons, often under difficult conditions. In some cases, one twin stayed with his or her biological parent(s), while the other twin was adopted away or was placed in the care of another family member. In the three original TRA studies published between 1937 and 1965, about 90% of the 75 pooled MZA pairs were born to "poor or situationally deprived parents."[94] It is likely that most twins experienced attachment-rupture trauma, emotional suffering, loneliness and neglect, abuse, and other types of

hardships. This was especially true for the late-separated children, and for children who spent time in an orphanage. As a group of French critics wrote, a more fitting name for genetic studies based on adoptees would be *the study of abandoned children*.[95] This evokes a different set of emotional responses than the more positive, trauma-obscuring terms "adopted children" or "twins reared apart." These terms fail to capture the rejection and abandonment trauma that many twins experienced, because they emphasize the fact that children were placed into adoptive homes, not on the psychologically damaging road they travelled to get there. If TRA studies were called "Studies of Abandoned Twins," people would view them very differently.

Research performed since the 1990s has shown that disturbed or ruptured caregiver-child attachment patterns can influence brain development during critical developmental periods.[96] As the authors of a 2015 review put it,

> "Environmental influences in infancy, particularly the quality of the caregiver–infant relationship and emotional interactions within this context, have been purported to shape neurological, psychological and social development and have potential long-term effects on psychological and emotional functioning.…Early deprivation of comfort and security has been found to have adverse sequelae on a broad range of domains, including neurological, psychological, emotional and physical development and functioning."[97]

These and other findings cast doubt upon the MISTRA researchers' claim that their conclusions apply (generalize) to the general (non-twin) population, as well as the their insistence that we must accept their conclusion that heredity plays a major role in causing behavioral differences, on the basis of how a few hundred abandoned twins performed on psychological tests.

The investigators performing the Finnish TRA study concluded that because their study did not represent an "ideal experimental situation," the "generalization of results may meet with some problems."[98] The MISTRA researchers easily and reasonably could have reached a similar conclusion in relation to their study, but their commitment to the promotion of genetic explanations of human behavior compelled them to apply their findings to most of the general population.

8. **Twins Were Not Placed Randomly into Adoptive Homes.** Three years prior to initiating the MISTRA, Bouchard wrote that "random assignment of cases to environments" was a "methodological necessity" in behavioral genetic research.[99] Bouchard later wrote that, although clearly "unethical," a "real [TRA study] experiment" would consist of twin pairs "literally separated at birth, randomly assigned to homes, and evaluated as adults prior to any social contact."[100] MZA twins, however, were not placed randomly, and here I will briefly describe two sources of bias that this introduced: selective placement bias, and restricted range bias.

Selective placement bias. A TRA study is both a twin study and an adoption study, because one or both twins were removed from their biological parents(s) and were placed into an adoptive or foster family, or with a relative. Kamin showed that MZAs and other adoptees were not randomly placed into available adoptive homes, but instead were *selectively* placed into such homes. *Selective placement* refers to adoption agencies' and others' practice or policy of placing adoptees into homes matching the socioeconomic (SES) and perceived genetic status of the birth (biological) parents. MZA pairs were not assigned at birth to different randomly selected adoptive homes spanning the entire socioeconomic spectrum, as would occur in a true scientific experiment. Researchers were unable to observe and test twins as they were growing up. Genetic interpretations of above-zero MZA group IQ and behavioral correlations, therefore, are based on the *assumption* that MZA pairs did not grow up in similar (correlated) environments. Although the MISTRA researchers claimed that selective placement bias was "modest" in their studies, non-random placements and correlated environments are the norm in TRA studies, especially when they are based on volunteer twins recruited through media appeals.

In his 1985 book *The Intelligence Men: Makers of the IQ Controversy*, psychologist Raymond Fancher described what he believed a "definitive" TRA study would look like:

> "A definitive study would have to employ twins who represent a genuinely random sample of the general population, and who have been *randomly* placed for adoption in a range of homes representative of the entire population. A definitive study would also have to demonstrate that its sample genuinely represents the full population of separated twins, and is not biased toward including only certain kinds of cases. Finally, in an ideal study all twins should have been *completely* separated from each other soon after birth, with no opportunity to communicate with each other or influence each other prior to their testing" (italics in original).[101]

None of the six published TRA studies came remotely close to meeting these standards, and I would replace Fancher's terms "definitive" and "ideal" with "scientifically valid." A scientifically valid study would also have to control or account for the environmental similarities experienced by even perfectly separated twins (see below). Otherwise, genetic interpretations of above-zero MZA behavioral correlations, even using the methods Fancher described, would remain contaminated by environmental factors, and MZA group correlations could be entirely explained by such factors. According to psychologist Robert Sternberg, a leading authority on human intelligence:

> "The method of separated identical twins…[has] limitations, such as the confounding variable that identical twins tend to be placed in similar, and hence correlated, environments, so that effects that may appear to be a result of genetic factors may, in fact, not be a result of such factors."[102]

And yet, the MISTRA results were based on the assumption that, as Bouchard put it, "twins were placed relatively randomly, with respect to trait-relevant environmental factors" (more on this assumption later).[103]

Range restriction bias. The restricted socioeconomic (SES) range of MZA family environments is another non-genetic factor biasing MZA correlations upwards. Since MZAs are usually placed in similar SES environments, on the basis of this restricted SES range alone we would expect them to correlate to a certain extent on IQ tests. Psychometrist Mike Stoolmiller showed in a 1999 publication that the restricted range of adoptive families inflates adoption study correlations that behavioral geneticists usually attribute to genetic influences. TRA studies are a type of adoption study, and Stoolmiller showed that these studies also "do not escape biases due to range restriction." He argued that "heritability estimates for related siblings adopted apart will also be seriously inflated by range restriction of SE [shared environment]."[104]

One can sympathize with researchers who for ethical, logistical, funding, and other reasons were unable to perform a proper study based on random assignment into homes reflecting the entire population, access to the complete population of MZA pairs through twin and population registers, and other aspects of a valid study. Nevertheless, we are not required to accept their conclusions simply because it was not possible for them to conduct such a study.

9. The Evidence Suggests that Most Studied MZA Pairs Were Only Partially Reared Apart. As the critics have shown, most MZA pairs in the original three TRA studies were only *partially* reared apart because most were separated late, grew up nearby to each other, and/or had substantial contact and a close relationship during important periods of their lives.[105] This also applies to the more recent TRA studies performed in Finland and Sweden, although the authors of these studies provided much less information than did the authors of the original three studies.[106]

For a number of reasons, it is unlikely that the MISTRA MZA pairs were any more "separated" than were the partially reared-apart pairs described in the earlier studies. Far from being separated at birth and living their entire pre-study lives not knowing that they had a twin, the MISTRA criteria stated that twins had "been separated by four years of age," and that they had "spent their formative years apart."[107] The researchers apparently decided that the twins' relationship after their (undefined) "formative years" was unimportant. (As an Internet search will show, the term "formative years" has several definitions and covers differing periods of development, depending on who defines it.) As Segal described the MISTRA age at separation:

> "The twins' mean age at separation was 218.21 days (standard deviation = 343.59) and ranged from 0.00 to 1,644 days. Expressed in years, the mean age at separation was 0.60 years (standard deviation = 0.94) and ranged from 0.00 to 4.50 years."[108]

Clearly, being "separated by four years of age," with the average age at separation being 0.6 years, is different from the impression given in many popular accounts that the MISTRA pairs were separated at birth. In his 2017 book *Behave: The Biology of Humans at Our Best and Worst*, Stanford University professor of biology and neurology Robert Sapolsky claimed that the "more than a hundred" MISTRA MZA pairs (actually 81) were "separated at birth." (Mukherjee also wrote that the MISTRA twins were separated at

birth.) Such statements help perpetuate the myth that TRA studies used only twin pairs who had been separated at birth, were reunited as strangers at the time of the study, and may have lived their entire (pre-study) lives not even knowing that they had a twin. In fact, this is rarely the case. Sapolsky saw the MISTRA as "the most wonderful, amazing, like, totally awesome thing ever in behavior genetics."[109] The best-selling Sapolsky and Mukherjee books of recent years have in common that their influential authors strongly endorsed the MISTRA findings, yet appeared to have been only superficially familiar with the original MISTRA publications.

In his 1981 book *Twins: An Investigation into the Strange Coincidences in the Lives of Separated Twins*, MISTRA-friendly journalist Peter Watson described a MISTRA reunited pair born in 1925, who had been separated at age two and a half (circa 1928), and who "had been seeing each other more or less regularly since 1946."[110] Another MISTRA pair, Dan Sivolella and Michael Meredith, had been separated at "six to eight weeks," and "first met up again when they were seventeen and then again at nineteen." They served together on the same U.S. Navy ship for a year, and later worked together in the same Indiana factory. They were 33 years old when studied by the MISTRA team, leaving them "plenty of opportunity to meet and influence one another" between the ages of 19 and 33.[111]

The story of another MISTRA MZA pair described by Watson is illuminating. Gladys Lloyd and Goldie Michael were 57 years old when they arrived in Minneapolis circa 1980, presumably after responding to a MISTRA media appeal:

> "In 1964, Gladys was married to a businessman who, behind her back, also wanted her twin as his mistress. He asked Goldie if he could buy a house for her, and a car, in return for her secret favours….Gladys found out and divorced the man. Since then the relationship between the sisters has been off and on. When they were both pregnant, they were close. They were also so alike then that Gladys's husband (another one) kissed Goldie by mistake as he came into the house one day from the office."[112]

According to the MISTRA criteria, Gladys and Goldie, in addition to another pair who "had been close friends for over thirty years,"[113] qualified as a pair of "reared-apart" MZ twins. Presumably (critics are not permitted access to the raw data), there are many more "separated" twins of this type in the MISTRA MZA sample.

Instead of making their information and data available for others to review, the MISTRA researchers devised the inadequate "contact time" formula, which assessed the amount of time that twins spent together before and after separation. Intimate relationships, however, are based on an ongoing association between people, which is not necessarily reflected by the amount of time they are in physical proximity. In addition, the concept made no distinction between the time spent together as infants versus time spent together as children or adults, and did not assess the quality of the contact. According to the MISTRA contact time formula, "Twins who met for a week at Christmas and for a week in the summer each year over a 10-year period are credited with 20 weeks of contact."[114] Instead of regarding twins who spent 14 days a year together over a 10-year span as having accrued 20 weeks of

contact time, it is far more accurate to say that these "reared-apart" twins had an ongoing 10-year relationship while growing up.

10. The MISTRA Twins Had Financial and Personal Incentives to Exaggerate Their Degrees of Separation and Behavioral Similarity. As Kamin noted in the 1970s, another key yet usually overlooked aspect of TRA studies is that the degree of MZA separation is based mainly on trusting the twins' accounts. Segal mentioned that "all travel and hotel expenses were covered by the project," plus a stipend of $260 (roughly $650 in 2018 U.S. dollars) for twins and their spouses to come to Minneapolis for the week, and "it became clear that having spouses, children, or friends travel with the twins provided an added incentive for the twins to participate."[115] It may have also provided an additional incentive for twins to exaggerate their degree of contact and behavioral similarity, or even to invent their separation.

Many MZA pairs came from England. Bouchard once commented that the fact that a large number of MISTRA twins came from England was "quite amazing when you think about it."[116] Indeed, 32.4% of the MISTRA twins were recruited from England, compared with 45.4% from the United States.[117] Because in the early 1980s the U.S. population was roughly five times larger than the English population, even assuming that recruitment procedures and the prevalence of MZA pairs were similar in both places, an unbiased sample would have been expected to contain about five times more American than English MZA pairs. The promise of an all-expense-paid transatlantic vacation for twins, some of whom could not have afforded one otherwise, may have provided motivation for the English pairs to describe themselves, and their degree of separation, in ways that would allow them to receive a free international vacation.

"Three Identical Strangers." A famous case of MZAs who exaggerated or invented their similarities is seen in the 2018 movie *Three Identical Strangers*. This movie chronicled the story of identical triplets Robert Shafran, Eddy Galland, and David Kellman, who were born in 1961 and were adopted away into three separate homes at six months of age as part of a secret and unethical unpublished study of separated twins, conducted by New York psychiatrist Peter Neubauer in the 1960s and 70s. The triplets grew up not knowing that they had identical twin brothers, until they discovered each other in 1980. Tragically, and probably related to the separation, Eddy Galland committed suicide in 1995.

In the early 1980s, the reunited triplets made the U.S. TV talk show circuit and became celebrities. They travelled to Minnesota and became MISTRA participants, and they were later featured in Segal's books. In their TV appearances the triplets were dressed alike and obviously staged some of their similar mannerisms, and they at times answered questions and finished each other's sentences in an obviously rehearsed way.

In the 1980s, the triplets attempted to profit from their fame on the basis of these claimed similarities. Although not mentioned in the movie, in a 1981 appearance on the U.S. *Today* show they revealed that they were enrolled in acting classes and hoped to star in a TV situation comedy. When that didn't happen, as seen in the movie they opened a New York

restaurant called "Triplets" that made over $1 million in its first year of operation. Decades later in *Three Identical Strangers,* Shafran confirmed, accompanied by a smile that suggested that he was understating the case, that he and his brothers emphasized their similarities and downplayed their differences. This was show business, not science.

Incentives for twins to lie about themselves to the researchers. Perhaps many MZA pairs were truthful, but there were incentives for them to exaggerate or to lie about their degree of separation, or to concoct "eerie" and "spooky" behavioral similarities between themselves. Motivations for MZAs to present themselves as more separated and more behaviorally similar than they actually were included:

- Financial, since they could potentially sell their stories to the media or to movie studios. Some MISTRA pairs hired agents, and like the triplets some appeared on television programs[118]
- The desire to be famous
- The desire to impress the researchers
- The desire to be the center of scientific attention and to feel important
- The opportunity to enjoy a fully paid vacation with stipend in Minneapolis for themselves and their spouses. Almost 50% of the MISTRA pairs came from overseas
- To feel that they had more of a bond with their co-twin

TRA researcher James Shields questioned the credibility and memories of the twins he studied:

> "Twins themselves will take delight in relating stories of their buying identical presents and perhaps asking an aunt to hide them in the same place. Or they claim, perhaps, to have changed their hair styles at the same time and to have decided independently to have their watches repaired before coming to London. They are said in many instances to come out with nearly the same remarks at the same time or to know what the other twin is thinking. Nearly all such stories have in common that they cannot be independently confirmed, and one sometimes suspects retrospective falsification of memory. Stories of twins falling ill at exactly the same time are not usually borne out by medical histories."[119]

Shields recalled with apparent amusement that reunited twins told him many lies, or at least embellished stories that "cannot be independently confirmed." As he put it, many pairs displayed a "retrospective falsification of memory." When twins told similar stories to the MISTRA team, they were (and continue to be) put forward by journalists, textbook authors, and the researchers themselves as "eerie" examples of the powerful influence of genetics. Shields attempted to verify many of the stories his twins told him. Did the MISTRA researchers attempt to verify their twins' stories as well?

11. The Key MISTRA Assumption That Above-Zero MZA Group Behavioral Correlations Are Caused Only by Genetic Factors Is Completely False. As stated in their 1990 *Science* publication, the researchers calculated heritability estimates "under the assumption of no [MZA] environmental similarity,"[120] and one of the main assumptions of their model was that "all resemblance between reared apart relatives is because of genetic factors." This implies that if genetic factors have no influence on behavior, and if there is no random or systematic error, the MZA group behavioral correlation would be zero (0.0). As early as 1983, in relation to the question of whether "environmental similarities explain the similarity in intelligence of identical twins reared apart," Bouchard answered with an emphatic "NO!"[121]

Many other genetically oriented researchers and commentators have endorsed this claim. MZA behavioral similarities, according to Mukherjee, "could have nothing to do with nurture; they could only reflect hereditary influences—nature."[122] In *Behave*, Sapolsky wrote that MZA "similarities in behavior probably reflect genetic influences."[123] According to Segal, "The MZA intraclass correlation of 0.50 for a personality trait directly estimates the heritability of the trait because MZA twins share only their genes."[124] And his 2004 book *Nature and Nurture: An Introduction to Human Behavioral Genetics*, Robert Plomin, the world's leading behavioral genetic researcher and also a co-investigator in the Swedish TRA study wrote, "If we report correlations from a study of identical twins reared apart, we implicitly have a very simple model that posits that resemblance is due solely to heredity." "How else," Plomin asked, "can we explain the twins' similarity?"[125]

The above statements are so obviously false, and so strongly go against basic common sense (simply being the same sex, the same age, and sharing a similar physical appearance will contribute to behavioral similarity), that it is hard to grasp that these influential authors actually believed what they were writing.

In addition to previously described biases such as the fact that most pairs were only partially reared apart, volunteer status, selective (non-random) placement, and range restriction, there are numerous non-genetic factors that contribute to MZA behavioral resemblance. Let's take a look at these additional influences now.

Age- and sex-confounds. Kamin identified age- and sex-effects as additional environmental confounds in TRA studies. A *confound* is an unforeseen or uncontrolled-for factor that threatens the validity of conclusions researchers draw from their studies. As a pair of genetically oriented researchers put it, "When you can't do experiments, you have to be very careful about something called *confounding*. Confounding is a pernicious problem that can make one thing look like it's causing something else when, in actuality, it's not."

Kamin found that similarity-biasing age-effects occurred in the earlier TRA studies when IQ tests were not properly standardized for age, meaning that MZA group correlations could be spuriously inflated due to "a defective age standardization of the I.Q. test." He made a similar point about potential sex-effects on IQ scores.[126] The MISTRA researchers were the first TRA investigators to recognize, as Kamin compelled them to recognize, that

"for most psychological, physiological, and medical variables there are substantial age and sex effects."[127] They devised a questionable and complicated statistical procedure, which they claimed corrected their correlations for these effects. In any case, we are about to see that age- and sex-effects on test scores are only a small component at least 28 potential environmental confounds at play in TRA studies.

MZA pairs, who are of course always the same age and sex, also share a common prenatal environment and a striking physical resemblance (e.g., height, attractiveness). Because in most societies men and women are socialized from birth to behave, think, and feel in differing gender-specific ways, the members of an MZA pair will behave more similarly for this reason alone. As a simple example, in Western societies *female-female* pairs are much more likely to resemble each other (correlate higher) for "lipstick-wearing behavior" than are *female-male* pairs. Examples of this type are usually overlooked by the popularizers of TRA studies.

There is also evidence of age effects on personality inventory scores, and evidence that personality can change over a person's lifespan. In a 1993 study of reared-together twins, McGue, Lykken and colleagues found that, over a ten year span, "some personality traits undergo normative and systematic changes."[128] And in a 2016 study, other researchers compared the personality test scores of the same person at ages 14 and 77. They found that

> "participants' and others' older-age personality characteristic ratings were moderately correlated with each other, and with other measures of personality and wellbeing, but correlations suggested no significant stability of any of the 6 characteristics or their underlying factor, dependability, over the 63-year interval."[129]

If personality inventory scores change over time in relation to developmental and other environmental influences, we would expect age-matched pairs to correlate on personality inventories simply because they are the same age.

Cohort effects. In addition to previously mentioned similarity biases, MZA pairs share many non-familial environmental influences in common, and are subject to non-genetic cohort effects. The *cohort effect* concept refers to similarities in age-matched people's behavior, preferences, beliefs, physical condition, and other characteristics that are caused not by heredity, but by experiencing stages of life at the same time in the same historical period and cultural milieu. In her 2015 autobiography, rock musician and leader of *The Pretenders* Chrissie Hynde (born in 1951) described the huge difference between her views and lifestyle as a young adult, versus those of her parents. Her behavior and beliefs were heavily influenced by the U.S. post-World War II "baby-boom" generation and the counterculture of the 1960s and 70s, whereas her conservative Republican Ohio parents were heavily influenced by the preceding Great Depression/World War II generation. Hynde and her parents were born at different times in different eras, and when growing up were influenced by very different ideas, peer groups, technologies, birth control options, financial conditions, wars, career opportunities for women, and countless other cohort influences.

Twins are of course born at the same time, and therefore are similarly exposed to cohort effects at the same points of their lives. Today in a typical big-city American coffeehouse, people reading the print edition of the daily newspaper are likely to be over 50 years old, not because of their genes, but because they grew up reading the newspaper this way. As behavioral geneticist Richard Rose pointed out in relation to the impact of cohort effects on MZA behavioral resemblance, "You're comparing individuals who grew up in the same epoch, whether they're related or not. If you asked strangers born on the same day about their political views, food preferences, athletic heroes, [and] clothing choices, you'd find lots of similarities. It has nothing to do with genetics."[130]

Even hypothetical MZA pairs (rarely found even in TRA studies) who were separated at birth, who never met each other, and who spent their entire lives not knowing that they had a twin, grow up experiencing many of the following 28 non-genetic prenatal and postnatal behavior-molding influences in common:

1. Prenatal (including common prenatal exposure to toxins and other influences)
2. Postnatal healthcare
3. Postnatal nutrition
4. Postnatal exposure to environmental toxins
5. Birth cohort (same age)
6. Gender cohort (same sex)
7. Developmental stage
8. Striking physical resemblance, including facial appearance and height
9. Adoptee status (with accompanying abandonment and attachment issues)
10. National
11. Regional
12. Political
13. Socioeconomic or class status
14. Ethnic/racial
15. Language
16. Religious (defined in part as "a cause, principle, or system of beliefs held to with ardor and faith")
17. Oppression, racism, discrimination, or privilege on the basis of common racial or national background, gender, SES status, language, religious beliefs, etc.
18. Climate/Weather
19. Shifting gender roles and increased career opportunities for women
20. Age at puberty onset or menarche
21. Diet/nutrition
22. Advertising and marketing campaigns
23. Exposure to the mass media, Internet, social media, etc.
24. Legal status of abortion
25. Birth control technology and availability
26. Selective placement status (adoption)
27. Teaching methods and technological advances
28. Exposure to similar music and lyrics

Susan Farber noted in *Identical Twins Reared Apart: A Reanalysis* that MZA pairs are "not so much similar to each other as they are similar to people of their eras and SES."[131] Although she was referring to twins' dental issues in this passage, this observation applies to IQ and behavioral similarity as well. Allow me to repeat and highlight this crucial point about MZA pairs' behavioral similarity and the meaning of TRA study findings:

MZA pairs are not so much similar to each other as they are similar to people of their eras and socioeconomic status

It therefore would be interesting to calculate psychological test-score correlations in a group of *genetically unrelated* pairs of age- and sex-like strangers, matched for SES and most of the above-listed environmental influences shared by MZAs.[132] It would help us understand the meaning and causes of MZA group behavioral resemblance.

TRA researchers' use of the word "apart" refers mainly to the *family* (rearing) environment, falsely implying that this is the only type of behavior-influencing environment that people can experience. This error is seen in Segal's claim that "MZA co-twins share all their genes but do not share their rearing environment," leading her to claim that "MZA co-twins have only their genes in common, so their observed similarities reflect their shared genes."[133] They actually have a lot more in common than their genes, since the rearing environment, as we have just seen, is only one aspect of the environments that twins experience throughout their lives. Travelers to some Middle Eastern countries, for example, will encounter the remarkable (non-genetic) behavioral similarity that most women wear head covering, even though almost all were "reared-apart" from each other in different family environments.

In a 1982 popular book on twins, journalist Kay Cassill described the MISTRA MZA pair Keith Heitzman and Jake Hellback, who grew up near New Orleans. Although the (apparently unbridgeable) "mighty Mississippi [river] divided these two physically, it could not separate their parallel lives. The welder from one side and the pump mechanic from the other found that they are both allergic to ragweed and dust. Both had done poorly in school. Both disliked sports and had cut their gym classes whenever they could. They are both addicted to candy. Their similarity of dress includes a penchant for wearing cowboy hats, which matches their parallel interest in guns and hunting." [134] As a critic pointed out, "Even if 'the mighty Mississippi divided' the twins, the fact that they both wear cowboy hats and like hunting is not that unusual for two [white] working-class men in the same region of Louisiana."[135] A similar point can be made about the famous "Jim Twins," two working-class white males who grew up in the same region of Ohio at the same time.

Clearly, there are numerous environmental influences unrelated to "common rearing" that contribute to MZA behavioral similarity and psychological test-score correlations. This means that the key MISTRA Step 3A/3B assumption that genetic factors are the only cause of above-zero MZA group psychological test-score correlations *is completely false*.

The researchers focused narrowly on statistical correlations, and ignored countless and obvious real-world examples of MZA environmental similarity, which they either assumed

do not exist, or counted as genetic effects. This helps illuminate the central fallacy of TRA studies: researchers claim that above-zero MZA group behavioral correlations are caused by shared genetic influences, when these correlations can be explained by cohort effects and other *non*-genetic influences and biases, plus researcher bias and error.

12. The Researchers' Model-Fitting Procedures Were Based on Assumptions That They Admitted "Are Likely Not to Hold." We have seen that the researchers based their conclusions in favor of strong genetic influences on IQ and other forms of behavior on the assumption that the MZA correlation "directly estimates heritability," on the assumption that environmental effects on MZA behavioral resemblance count as genetic effects (see the discussion in the next section), and on the results of model-fitting analyses. We have also seen that heritability estimates are generated by these procedures (Step 3A).

An assumption of the MISTRA model-fitting analyses, as described in a 1989 publication by McGue and Bouchard, was that the "phenotype (i.e., character of interest) can be expressed as a simple additive function of a genetic effect *(G)* and an environmental effect *(E)*." Similar to Segal's previously discussed description, "Additional assumptions" of the MISTRA model included,

> "**(a)** All resemblance between reared apart relatives is because of genetic factors, **(b)** there is no assortative mating, **(c)** all genetic effects are additive (i.e., there is no dominance or epistasis), and **(d)** genetic and environmental effects are independent (i.e., there no genotype-environment co-variance) and combine additively in the determination of the phenotype (i.e., there is no genotype by environment interaction)" (bolding added).[136]

Although rarely if ever reported in the scientific or popular literature, or mentioned by Segal in her books, in this 1989 publication McGue and Bouchard recognized that "several of these assumptions are *likely not to hold* for cognitive abilities" (italics added).[137] This assessment clearly applies to personality and other behavioral characteristics as well.

Critics have argued since the 1970s that model-fitting analyses are based on implausible assumptions, and even pioneers of the technique admitted, amazingly, that the "elementary model…makes a number of very strong assumptions which may not be generally true."[138] Let us look briefly at the assumptions underlying the MISTRA model.

Assumption A. We saw earlier that the MISTRA "Assumption a" that "all resemblance between reared apart relatives is because of genetic factors" is completely false, since even perfectly separated MZA pairs experience many common environmental influences. The model also assumed that "twins were placed relatively randomly, with respect to trait-relevant environmental factors."[139] Researchers can make this assumption only if they are able to identify the *specific and exclusive* environmental influences that contribute to the behavior in question, and then determine that twins did not experience such influences, or that they experienced them to the same degree. Bouchard, however, believed that we do not know what "trait-relevant" influences might cause differences in IQ and personality. "In

spite of years of concerted effort by psychologists," he wrote in 1997, "there is very little knowledge of the trait-relevant environments that influence IQ…and ordinary personality traits."[140] Bouchard, therefore, could not logically assume that MZAs did not experience "trait-relevant" environmental factors for IQ and personality, since he had no idea what these trait relevant factors might be.

Assumption B. The MISTRA "Assumption b" was that "there is no assortative mating." According to Wikipedia, *assortative mating* refers to a "mating pattern and a form of sexual selection in which individuals with similar phenotypes mate with one another more frequently than would be expected under a random mating pattern." Of course, if humans carry no genetic variants that cause differences in IQ and other behaviors, then there is no assortative mating bias because mating patterns would have no direct genetic influence on these behaviors. If we assume that humans do carry such variants, then it is clear that assortative (non-random) mating is a standard human practice for many types of behavior. As the critical behavioral geneticist Jerry Hirsch (1922-2008) put it, "Our species mates assortatively, not randomly, e.g., my wife and I met and married in Paris as American students at the Sorbonne…analogous stories are true of countless colleagues. Yet, heritability estimation assumes…random mating."[141]

Assumption C. The MISTRA "Assumption c" that there is no dominance or epistasis may also be false. Epistasis occurs when the effect of one gene depends on the presence of one or more "modifier genes." Some research suggests that epistasis is "strong" and "pervasive" in mice and rats,[142] and is reported to influence human characteristics as well.[143]

Assumption D. The MISTRA model's "Assumption d" position that "genetic and environmental effects are independent" and "combine additively" is false, since genetic and environmental effects are bidirectional. As psychologists Douglas Wahlsten and Gilbert Gottlieb wrote in 1997, the behavioral genetic claim that "heredity and environment are additive, separately acting causes" is "biologically unrealistic in view of all that is known today about the control of gene action and the interdependence of genetic and environmental effects."[144]

The MISTRA researchers hoped that their model's false assumptions would "offset each other" in favor of genetics. While recognizing in 1989 that the main assumptions of their model were "likely not to hold," in 2007 the MISTRA researchers wrote that their model-fitting assumptions "are generally oversimplifications of the actual situation, and their violation can introduce systematic distortions in the estimates." They claimed, however, that "several combinations of violations of assumptions can act to offset each other."[145] Such genetically biased speculation—that numerous false assumptions somehow all cancel each other out in favor of genetics—allowed them to conclude what they and their funding sources needed and wanted to conclude, namely, that above-zero MZA group correlations, either standing alone or worked into model-fitting statistical analyses, showed that genetic factors have a strong and pervasive influence on most areas of human behavior.

Clearly, a set of findings and conclusions that researchers correctly recognize are based on assumptions that "are likely not to hold," which they hope will "offset each other" in favor of genetics, is not a valid set of findings and conclusions. Sociologist Howard Taylor showed in his 1980 book *The IQ Game: A Methodological Inquiry into the Heredity-Environment Controversy* that TRA and other types of behavioral genetic studies are based on a "string of flimsy and implausible assumptions."[146] However, one would not know this by relying on mainstream accounts and textbooks, whose authors rarely mention the "flimsy and implausible" assumptions upon which the MISTRA findings were based.

13. The MISTRA Assumption That Environmental Influences Shared by MZA Pairs Should Be Counted as Shared Genetic Influences is Fallacious.

Amazingly, Bouchard and colleagues based the MISTRA conclusions on the claim that environmental influences on MZA group behavioral resemblance should be counted as *genetic* influences. In their 1990 *Science* article, they wrote that one of the three "implications" of their genetic "findings," and of behavioral genetic findings in general, was that

> "MZA twins are so similar in psychological traits because their identical genomes make it probable that their effective environments are similar.…It is a plausible conjecture that a key mechanism by which the genes affect the mind is indirect, and that genetic differences have an important role in determining the effective psychological environment of the developing child."[147]

The above statement is not an implication of the researchers' findings; rather, it is an assumption upon which they *based* their findings. Bouchard and colleagues believed that MZA behavioral resemblance caused by the impact of environmental influences "is counted as a genetic influence" because MZA pairs' "identical genomes" cause them to create more similar environments for themselves.[148] As they wrote in their 1990 *Science* article, the "radical environmentalist" belief that the "the proximal cause of most psychological variance probably involves learning through experience" is probably correct. However, they transformed these environmental causes into genetic causes on the basis of their claim that exposure to similar environments is "guided by the steady pressure of the genome."[149]

Looking back in 2016, Bouchard wrote,

> "Our interpretation of the results of MISTRA was very straightforward. We expected that with regard to psychological traits, monozygotic twins reared apart were similar because their effective environments were similar. This was because their environments were self-selected and that selection was guided by their genotype."[150]

This statement could be interpreted as unintentionally overturning the results of the entire study. Bouchard wrote that he and his colleagues expected substantial MZA group correlations on psychological tests because the twins' *environments* were similar. However, in the case of IQ, even if shared behavioral genes cause MZAs to behave more similarly and to create similar IQ-relevant environments for themselves, there is no reason to believe

that these genes have anything to do with IQ or "general intelligence." Using a relevant hypothetical example from medicine, if tobacco smoking behavior is under complete genetic control, it doesn't mean that lung cancer is caused by cancer genes. Based on Bouchard's logic, we should conclude that lung cancer caused by smoking is completely under genetic control because lung damage caused by the carcinogenic effects of tobacco tar is "self-selected," and that "selection was guided by" a tobacco smoker's genotype.

Two seemingly counterposed previously cited MISTRA statements are as follows:

"Monozygotic twins reared apart were similar because their effective environments were similar"

versus

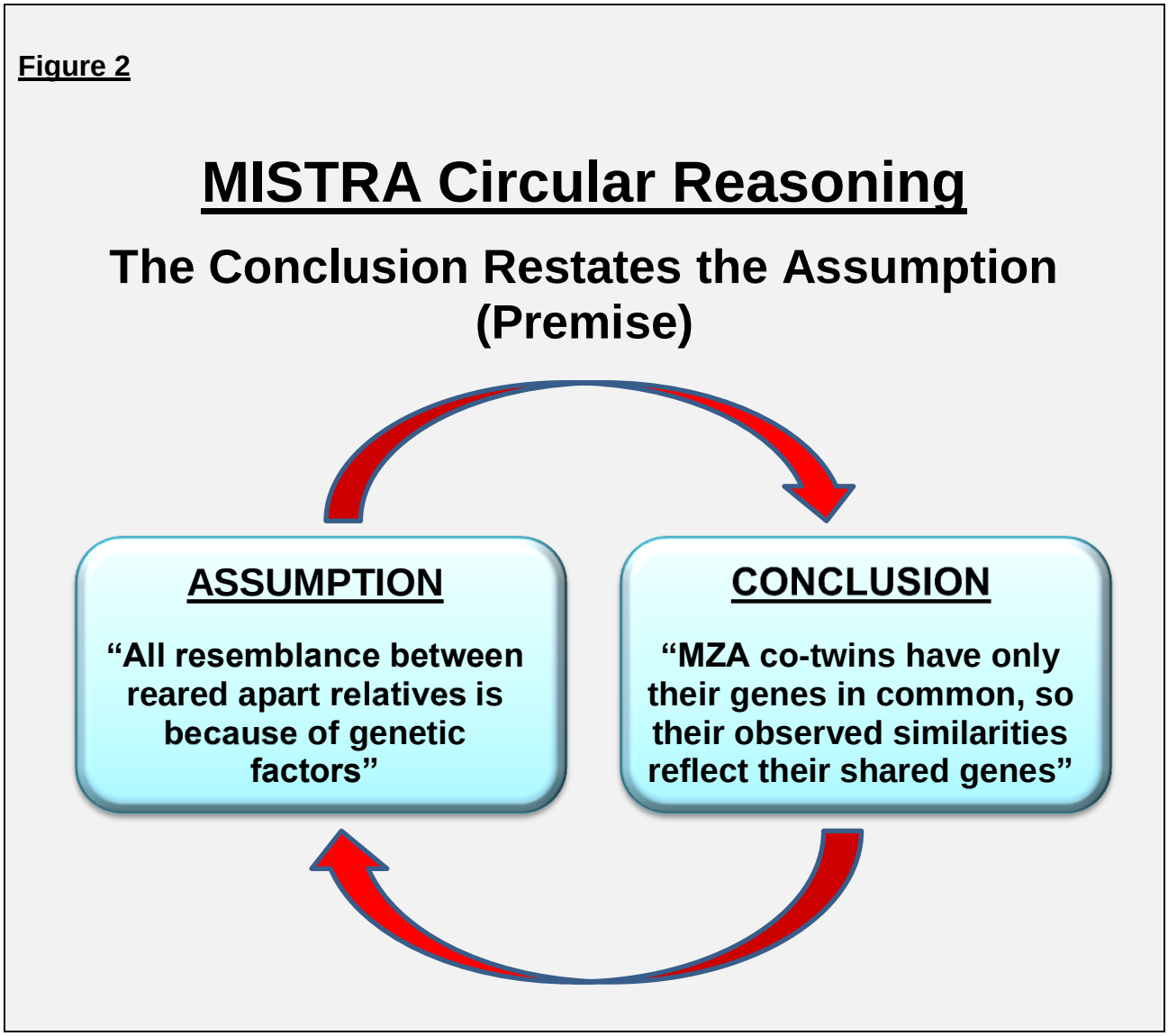
"All resemblance between reared apart relatives is because of genetic factors"

We have seen that the researchers attempted to reconcile these contradictory statements by deciding to count environmental effects as genetic effects. This confusion is found in the 1990 MISTRA *Science* article, where we saw that they based their heritability estimates on the assumption of "no [MZA] environmental similarity," and then concluded that "MZA twins are so similar in psychological traits because their identical genomes make it probable that their effective environments are similar."[151]

Circular reasoning. As I have shown elsewhere in relation to twin method MZT-DZT comparisons, the MISTRA argument that MZA twins create similar environments for themselves because they are genetically similar is a fallacious one.[152] The argument is based on circular reasoning, which is the logical fallacy of assuming the very thing that one is attempting to determine, where the premise depends on or is equivalent to the conclusion.

In *Born Together—Reared Apart*, Segal wrote, "The bottom line from our data was that growing up together does not make family members alike. Instead, our findings showed that personality similarity between relatives seems to come mostly from their shared genes."[153] Aside from the fact that Segal overlooked the numerous non-familial environmental influences that "make family members alike," the claim that "personality similarity between relatives seems to come mostly from their shared genes" is both a MISTRA finding and a MISTRA assumption. As seen in Figure 2, the MISTRA researchers concluded that behavioral resemblance among relatives is caused by their shared genes *based on the assumption* that behavioral resemblance among relatives is caused by their shared genes. It appears that behavioral genetic defenses of twin studies—whether using the "reared-apart" or "reared-together" design—are based on obviously fallacious arguments that have been overlooked by academic fields for decades.

Figure 2



As seen in Figure 2, the MISTRA position that genetic factors cause MZA group behavioral resemblance is both an assumption (premise) and a conclusion of the study. This is the circular fallacy of arguing that "X is true because Y is true; Y is true because X is true." Bouchard, Segal and their colleagues, therefore, referred to their assumption in support of their conclusion, and then referred back to their conclusion in support of their assumption, in a circular loop of faulty reasoning.

Do twins create their environments? Bouchard, Segal and colleagues wrote in their 1990 *Science* article that environmental effects count as genetic effects in part because "infants with different temperaments elicit different parenting responses."[154] This claim portrays twins (and children in general) as behaving according to an inherited behavioral

blueprint, but implies that adoptive parents (and other adults) are easily able to change *their* behavior and treatment in response to the twins' behavior—in effect being flexible enough to let twins "create their own environments." However, parents are people too, and the MISTRA "twins create their environments" scenario ignores the fact that, according to this scenario, adults' "parenting response" behavior must be far more unchangeable than the twins' supposed "parent response eliciting" behavior. This is because, in addition to the parents' presumed "pronounced" genetic predispositions, they have experienced decades of peer, family, religious, cohort, and other behavior-molding influences.

In the MISTRA, factors that influenced above-zero MZA group behavioral correlations were subsequently considered—by definition—to be genetic factors, thereby creating a genetic "heads I win, tails you lose" type of study that guaranteed that genetic explanations would prevail.

14. **The MISTRA Claim/Assumption That the "MZA Correlation Directly Estimates Heritability" Is False.** The researchers' Step 3B claim that the MZA correlation "directly estimates heritability" was based on what we have seen is the false assumption that only genetic factors can account for above-zero MZA group correlations. And yet, claims of strong genetic influences on IQ are based on this claim.

    According to Bouchard, "The correlations for monozygotic twins reared apart directly estimate the heritability of the trait."[155] And Segal wrote in *Born Together—Reared Apart* that the "MZA intraclass correlation directly estimates broad heritability because MZA twins share all their genes but do not share their rearing environment."[156] For example, if the MZA group IQ correlation is .70, IQ heritability would be estimated at .7 (70%). We have seen, however, that there are at least six major problems with this position:

    - Most studied MZA pairs were only *partially* reared-apart
    - The MZA sample was biased in favor of behaviorally similar pairs
    - Even perfectly separated MZA pairs experience many non-familial environmental influences in common
    - Environmental influences should not be counted as genetic influences
    - In the area of IQ, the researchers omitted their DZA group correlations (they bypassed Steps 2 and 3A)
    - There are major problems with the heritability concept itself

15. **The Researchers Failed to Publish Their Full-Sample Control Group DZA IQ Correlations.** We have seen that the researchers completely removed their DZA group results from the 1990 *Science* IQ study, even though they published full-sample DZA correlations for most non-IQ behavioral characteristics in other publications. Although DZA pairs constituted the MISTRA designated control group, the MISTRA full-sample DZA IQ correlations remain unpublished to this day. In the late 1990s the researchers published *sub*sample IQ or "general cognitive ability" DZA group correlations, in the context of assessing possible genetic influences on "right-wing authoritarianism" and "ego development."[157]

Although the investigators have never published their full-sample WAIS, Raven/Mill-Hill, or First Principal Component IQ DZA group correlations, they did publish their full-sample DZA correlations for "special mental abilities" (verbal, perceptual, spatial, and memory tasks). When DZA correlations in this and other areas did not fit genetic predictions, they explained them in ways that were consistent with genetic theories. If the DZA correlation was one-half of the MZA correlation, they concluded that this result fit genetic expectations. If the DZA correlation was low versus the MZA correlation, or even negative, they had "flagged a possibly emergenic trait," which provided a genetic explanation for a DZA correlation not significantly above zero.[158] If the DZA correlation was not significantly different from or even higher than the MZA correlation, the researchers assumed that the culprit was "low sample size," "assortative mating," "statistical variation," or "sampling variability."[159] Consistent with the MISTRA researchers' genetic biases, when commenting on a DZA full-sample or subsample correlation that they chose to publish, they had a genetic explanation for it regardless of its value.[160]

A pair of 1990 MISTRA publications discordant for reporting DZA group correlations. Some of these anomalous DZA correlations, in this case special mental abilities correlations, are found in a 1990 MISTRA paper published in the little-known twin research journal *Acta Geneticae Medicae et Gemellologiae*, then published in Rome (now continued as *Twin Research and Human Genetics*). While noting that some of their DZA group correlations were difficult to explain on genetic grounds, Bouchard and colleagues stuck to their position that "a genetic component was needed to statistically account for the twin data," while making a rare MISTRA concession that "alternative explanations are obviously possible and are being explored."[161]

As we have seen, Bouchard and colleagues began their 1990 *Science* publication with the statement that "monozygotic and dizygotic twins who were separated early in life and reared apart (MZA and DZA twin pairs) are a fascinating experiment of nature. They also provide the simplest and most powerful method for disentangling the influence of environmental and genetic factors on human characteristics." Yet stunningly, unlike the 1990 MISTRA *Acta* article, the *Science* article contained no DZA correlations or model-fitting analyses of any kind. Bouchard and colleagues claimed that they omitted their control group DZA correlations "due to space limitations and the smaller size of the DZA sample (30 sets)."[162] It is almost as if they forgot to remove the 1990 *Science* article opening paragraph after deciding against publishing their DZA group correlations (possibly reflecting a last-minute decision to omit the DZA correlations from the published version of their article). The *New York Times* writer who reported on the study then wrote, erroneously, that the researchers arrived at their conclusions "by comparing intelligence test results of identical and fraternal twins who were brought up separately."[163] (It is possible that the writer was given a pre-publication version that reported and compared the MZA and DZA correlations.)

These two 1990 MISTRA publications illustrate an additional bias in the process of deciding where to submit and publish research results. Findings and analyses published in obscure specialty academic journals such as *Acta Geneticae Medicae et Gemellologiae* go largely unnoticed, and were even more unnoticed in the pre-Internet/pdf era, whereas the

worlds of academics and journalism pay a great deal of attention to articles published in *Science* and other top journals.

More on the missing DZA group IQ correlations. In her *Born Together—Reared Apart* chapter on the MISTRA IQ and personality studies published in the period 1988-1990, Segal wrote that the 1990 *Science* paper "reported IQ analyses for forty-eight MZA twin pairs for whom we had processed data; the [unpublished] DZA twin sample was still modest (thirty pairs)."[164] Although she believed that "DZ twins reared apart (DZA) constitute and important control group,"[165] Segal provided no other explanation for the researchers' decision to omit their DZA control group correlations, and then went on to discuss the study and its finding of major genetic influences on IQ based on Step 3B and Step 4.

Although the researchers claimed that their 1990 DZA sample was "modest" in size, and that there was not enough "space" to publish their DZA correlations, a 1990 MISTRA study of personality contained full-sample DZA group personality correlations based on only *26* pairs.[166]

In a 1998 academic journal article, Bouchard wrote that although the DZA group correlation "is reported as a control variable," the full-sample "MISTRA [DZA IQ] correlations have not yet been fully analyzed" because he and his colleagues were "awaiting completion of the study before conducting a full analysis."[167] Bouchard would have had us believe—almost two decades into the study—that he had not published his full-sample MISTRA DZA group IQ correlations because they had not been "fully analyzed."

Five years earlier, Bouchard had reported that the concurrently running "Swedish Adoption/Twin Study on Aging" (SATSA) TRA study was the only such study to have reported full-sample DZA IQ correlations. This was true, but only because Bouchard had decided against publishing the *MISTRA* full-sample DZA group IQ correlations.[168]

All this occurred despite the fact that IQ was the most important focus area of the study, that "the study of IQ is paradigmatic of human behavior genetic research," and that the MISTRA researchers had designated DZA twins as their study's control group.[169] When Kamin wrote to Bouchard in 1997 asking him to supply the complete DZA IQ data, Bouchard refused to do so, replying, "I can't pass on the IQ results for our MZAs or DZAs because I have not published them yet. Indeed, I have not even calculated them."[170] In other words, in this "paradigmatic" area of behavioral genetics, the researchers ***failed to publish or share the full-sample IQ correlations produced by their own designated DZA control group.*** In the next section I will show that this occurred, most likely, because the results failed to show a genetic influence on IQ.

16. The Published *Near* Full-Sample MZA and DZA IQ Correlations Did not Differ at Statistically Significant Levels, Supporting the Critics' Conclusion That the Study Found No Evidence in Support of Genetic Influences on IQ. We have seen that because MZA pairs are more similar to each other genetically than are DZA pairs (theoretically 100% versus an average 50%), genetic theories predict that the mean MZA

group behavioral correlation will be higher than the corresponding DZA group correlation at a statistically significant level (Step 2). If not, we can conclude that non-genetic factors alone were responsible for raising both correlations above zero, since MZAs' greater genetic resemblance did not lead to their greater behavioral resemblance versus DZAs.

The MISTRA genetic model predicted that the MZA group correlation will be high, and that the DZA group correlation will be about one-half the size of the MZA correlation, or possibly lower. As Bouchard wrote in a 1994 publication, "If MZA twins were very much alike and DZA twins showed a level of similarity less than half the MZA twins, this would be evidence for configural genetic influence."[171] The MISTRA formula was *r*DZA = .5h^2, meaning that the DZA correlation was expected to be one-half (.5) of the heritability estimate.[172] Because the MZA correlation and the heritability estimate were assumed to be the same value ("the MZA correction directly estimates heritability"), this meant that, according to the MISTRA genetic models, the MZA correlation was expected to be high, and was also expected to be significantly higher and about twice as large as the DZA correlation.

Earlier I mentioned the hypothetical example of an Argentinian TRA study of spoken language. In such a study the MZA and DZA correlations for speaking Spanish would be very high and very similar, which would lead to the conclusion that the study found no evidence that genetic factors influence the language that a person speaks. Turning to IQ, Nancy Pedersen, Plomin and their colleagues observed in relation to their 1992 Swedish TRA study, "When MZ correlations are not greater than DZ correlations, twin similarity may reflect correlated environments rather than genetic similarity."[173] The principle that applies to spoken language behavior, as we see, applies to IQ-score behavior as well.

Nancy Segal and others on the importance of the MZA-DZA comparison (Step 2). Although we saw earlier that the researchers arbitrarily bypassed the Step 2 direct comparison between their MZA and DZA group IQ correlations (see Figure 1), Segal confirmed in *Born Together—Reared Apart* that determining whether the MZA group correlation is higher than the corresponding DZA correlation is "an important first step" in demonstrating "whether or not" genetic factors influence a behavior:

> "The simple comparison of the MZ (or MZA) and DZ (or DZA) intraclass correlations is *an important first step* in behavioral-genetic analysis because this demonstrates *whether or not* there is genetic influence on the trait" (italics added).[174]

Elsewhere in the book Segal wrote,

> "Genetic effects are shown *if* the correlation for MZ or MZA twins exceeds the correlation for DZ or DZA twins" (italics added).[175]

In Plomin and colleagues' 2013 edition of their textbook *Behavioral Genetics*, genetic researcher Sean Purcell described the "the important first question" that researchers "must ask" when performing a model-fitting analysis:

> "Simple comparisons between twin correlations can indicate whether genetic influences are important for a trait. This is the important first question that any quantitative genetic analysis must ask."[176]

And in a 1990 publication, Plomin and colleagues noted the limitations of model-fitting analyses when the MZ correlation is not significantly higher than the DZ correlation:

> "We should not stand too much in awe of model fitting or allow it to obfuscate the basic simplicity of most behavioral genetic designs. For example, the twin design estimates genetic influence on the basis of the difference between MZ and DZ correlations. *If the MZ correlation does not exceed the DZ correlation for a particular trait, there is no genetic influence* (unless assortative mating approaches unity), and model-fitting analyses must come to that conclusion or there is something wrong with the model" (italics added).[177]

As Plomin stressed here and in the Second and Third Editions of his textbook *Behavioral Genetics*, a model-fitting analysis finding genetic influence would be wrong if the MZT correlation does not (significantly) exceed the DZT correlation, a conclusion that applies to MZA-DZA comparisons as well (Step 2).[178] If the MZ correlation is not significantly higher than the DZ correlation, he wrote that "there is no genetic influence" on the trait.

The MISTRA MZA and DZA IQ correlations did not differ at a statistically significant level. As it turns out, the MISTRA IQ correlations failed to pass the Step 2 test described by Bouchard and colleagues in 1986, Plomin and colleagues in 1990, Pedersen and colleagues in 1992, Segal in 2012, and Purcell in 2013.

As we have seen, the final 2000 MISTRA sample consisted of 81 MZA and 56 DZA pairs. Based on a *near* full-sample 2007 MISTRA study of 74 MZA and 52 DZA pairs by Wendy Johnson, Bouchard, Segal and others, there is no statistically significant difference between the MZA and DZA group correlations for either the Wechsler (WAIS) or the Raven's Progressive Matrices tests, which were the main IQ tests used in the MISTRA. The DZA sample of 52 pairs contained 14-18 opposite-sex pairs.[179] Because same-sex DZA pairs experience more similar environments than experienced by opposite-sex pairs, the opposite-sex pairs probably lowered these DZA correlations when compared with MZAs, who are always the same sex.

As seen in Table 2, the near full-sample MISTRA Wechsler full-scale IQ correlations, as reported by Segal in 2012 almost as an afterthought, were MZA group = .62, versus DZA group = .50.[180] These Wechsler correlations were based on the unpublished 2007 figures given to Segal by Bouchard in 2009. The near full-sample MISTRA Raven IQ correlations were MZA =.55, versus DZA = .42.[181] I am unaware of any publication reporting the MISTRA "First Principal Component of Special Mental Abilities" DZA group IQ correlations, another important omission that has been overlooked by mainstream commentators, and authoritative authors and researchers, for decades.

Table 2

Near Full-Sample MISTRA IQ Correlations

MZA vs. DZA Twin Pairs

	MZA Pairs (Experimental Group)	**DZA Pairs** (Control Group)	**Probability Value (p)**
Wechsler IQ (WAIS)	*r* = **.62** (74 pairs)	*r* = .**50** (52 pairs)	**p = .17** Not statistically significant at the .05 level
Raven's Progressive Matrices IQ	*r* = **.55** (74 pairs)	*r* = **.42** (52 pairs)	**p = .18** Not statistically significant at the .05 level

The final MISTRA full sample consisted of 81 MZA and 56 DZA pairs.

Based on calculations made at the VassarStats website. *r* = intraclass correlation; p = one-tailed probability; MZA = monozygotic twins reared apart; DZA = dizygotic twins reared apart; MISTRA = Minnesota Study of Twins Reared Apart; WAIS = Wechsler Adult Intelligence Scale. Sources: **Wechsler (WAIS)** correlations are from Segal, N. L., (2012), *Born Together—Reared Apart,* Cambridge, MA: Harvard University Press, p. 286, based on the number of pairs reported on p. 284; **Raven** correlations are from Johnson et al., (2007), *Intelligence, 35,* 542-562, p. 552, Table 3, based on the number of pairs reported on p. 545. The WAIS and the Raven were the two main IQ tests used in the MISTRA. The DZA sample contained 14-18 opposite-sex pairs.

As we see in Table 2, statistical tests show that both the MISTRA Wechsler and the MISTRA Raven MZA versus DZA "important first step" correlation comparisons failed to differ at the conventional .05 level of statistical significance. This leads to the conclusion that, for each test, the MZA and DZA group IQ correlations are the same, because it is assumed that the differences occurred by chance (the null hypothesis stating that the correlations do not differ is not rejected).

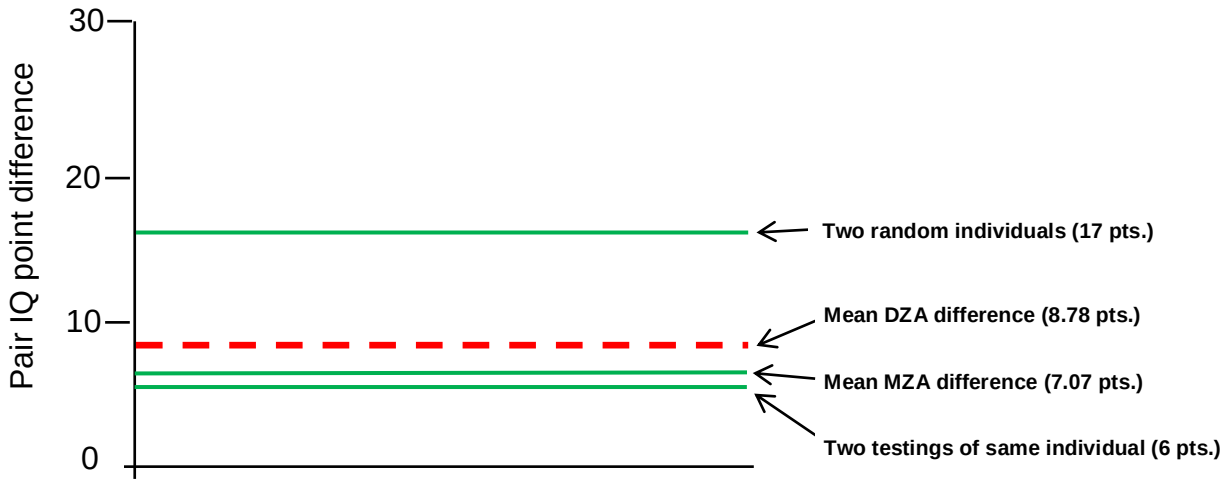
In her 2017 book *Twin Mythconceptions: False Beliefs, Fables, and Facts About Twins*, while citing no sources or data in support of this statement, Segal wrote, "If…MZA twins are more alike in intelligence than DZA twins (*which they are*), this tells us that genes affect intellectual development" (italics added).[182] On the contrary, we see in Table 2 that according to basic methods of statistical significance testing between two mean sample correlations, MZA twins *are not* "more alike in intelligence" than DZA twins.

Contrary to the way the MISTRA results are usually discussed in the scientific literature (including textbooks) and in the popular press, for this reason alone we can safely conclude that the study failed to identify a genetic influence on IQ. Most likely, Bouchard did not publish, share, make available, "fully analyze," or "calculate" his full-sample DZA IQ correlations because they led to the undesired conclusion that his study failed to find genetic influences on IQ scores (general intelligence). ***In other words, the potentially conclusion-overturning MISTRA full-sample control group DZA IQ correlations were, and remain, suppressed.***

Mean difference scores, reported and non-reported. If we look at MZA and DZA IQ scores in terms of the average absolute mean differences between twins' scores, we see in Figure 3 that, as Segal reported in 2012, the MZA group mean difference was 7.07, and the DZA mean difference was 8.78.[183] These can be compared with the expected 17-point difference between two randomly selected unrelated people.[184]

Figure 3

1990 MISTRA *Science* Article Within-Pair Mean IQ Differences: Reported and Missing



Adapted from the Bouchard et al. 1990 *Science* article, Figure 1, p. 227. The solid **green** horizontal lines are adapted from this figure. The dashed **red** horizontal line represents the missing DZA group IQ mean difference. Pts. = mean IQ point difference. MZA and DZA mean difference scores were reported in Segal, 2012, *Born Together—Reared Apart*, pp. 286-287. Mean difference scores for random individuals, and for two testings of the same individual, as reported in Plomin, R., & DeFries, J. C., (1980), *Intelligence, 4,* 15-24, p. 23 (this also was the source cited in the 1990 MISTRA *Science* article). Information on pair contact reported in the MISTRA figure is omitted here.

Figure 3 is adapted from the 1990 MISTRA *Science* article's "Figure 1."[185] In the context of assessing pair mean difference scores as a function of the degree of "pair contact," Bouchard and colleagues drew lines comparing the MZA group IQ mean difference, the expected IQ difference between "two testings of [the] same individual," and the expected IQ difference in a group consisting of "two random individuals." These three lines are reproduced in Figure 3 as solid green lines. The first two mean difference lines were close together in the 1990 figure, and the unrelated-pair difference was far above these lines. This led Bouchard and colleagues to comment that we can "observe a wide range of differences" in the figure, since MZA pair IQ scores were much more similar than were scores among randomly selected unrelated pairs.[186] The dashed red line that I have added in Figure 3 represents the DZA control group IQ mean difference as reported by Segal in 2012, which was omitted from the 1990 *Science* figure. As we see, the DZA mean

difference line is very close to the MZA line. In fact, because it is unlikely that the 1990 MZA and DZA IQ means differed from each other at a statistically significant level, the MZA and DZA mean difference lines would have been the same line.[187] Once again, by failing to include the control group DZA results, Bouchard and colleagues provided the readers of their *Science* article with an incomplete and therefore distorted account of the study's actual findings.

***

In his 2003 book *The Agile Gene: How Nature Turns on Nurture*, genetically oriented science writer and MISTRA popularizer Matt Ridley emphasized the fact that the MISTRA was the first TRA study to use a DZA control group. To answer TRA study critics, he wrote,

> "Bouchard…set out to find fraternal (dizygotic) twins reared apart [DZAs]. These were people who [like MZAs] shared a womb as well as a western upbringing. If his critics are right, then they too should show remarkable similarities of the mind. Do they?"[188]

Ridley, who believed that the MISTRA results helped "twin studies silence their critics," answered this question in the negative by listing several MISTRA-studied behaviors that showed a significantly higher MZA versus DZA correlational difference. Predictably, however, he said nothing about IQ scores in this context. As seen in Table 2 and in Figure 3, the IQ results suggest that the Minnesota MZA and DZA pairs did indeed show "remarkable similarities of the mind," providing yet another reason to reject genetic interpretations of the MISTRA IQ results.

17. **The Researchers' Computer Software Program Was Designed to Favor Genetic Explanations.** In the 2007 Johnson et al. MISTRA publication, Bouchard, Segal, and colleagues revealed how they worked their assumption that genetic factors strongly influence cognitive abilities into their statistical analyses of the data. Clearly referring to their MZA and DZA pairs they wrote, in a footnote,

> "For 4 tests, DZ correlations actually exceeded MZ correlations, a situation we attribute to sampling variability. In such situations, [the] Mx [software program] gives greater weight to the larger MZ than DZ sample, providing estimates of genetic influence based primarily on the MZ correlations."[189]

Note their decision to "attribute" high DZA group correlations to "sampling variability," and not to conclude that the results failed satisfy the Step 2 requirement that the MZA group correlation must be significantly higher than the DZA correlation—once again ignoring the correlations of their designated control group. They simply assumed that higher DZA correlations, unexplainable by genetic theories, were the result of sampling variability, and then used their genetically biased software to greatly diminish the impact of the DZA correlations and to base genetic estimates mainly on the MZA group correlations, which they assumed directly estimate heritability. This is seen in their 2007 table of MZA

and DZA test-score correlations where, as one example, Wechsler Information subscale correlations of MZA = .57, DZA = .64 still produced an "estimated genetic influence" (heritability) figure of .65![190] Because this comparison (and many others) failed to pass the Step 2 requirement of finding a significantly higher MZA versus DZA correlation, the researchers should have concluded that they found *no* genetic influence on the Information subscale and other measures of cognitive ability. Psychologist Ken Richardson has written that behavioral genetic researchers make "certain assumptions about 'what to expect' in the patterns of scores, and adjusted their analytical equations accordingly: not surprisingly, that pattern emerges!"[191]

As the saying goes, garbage in, garbage out.

18. **Personality Inventories (Tests) and the "Personality" Concept Are Controversial.** Journalists and the authors of influential works usually ignore the fact that Juel-Nielsen, in his 1965 register-based study, found "marked intra-pair" personality differences in all 12 MZA pairs he studied:

> "In all 12 pairs there were marked intra-pair *differences* in that part of the personality governing immediate psychological interaction and ordinary human intercourse. . . .The twins behaved, on the whole, very differently, especially in their cooperation, and in their form of and need for contact. Corresponding with these observations, the twins gave, as a rule, expression to very different attitudes to life, and very divergent views on general culture, religion and social problems. Their fields of interest, too, were very different. . . .Those twins who had children treated, on the whole, their children differently, and their ideas on upbringing were, as often as not, diametrically opposed. Characterologically, the twins presented differences in their ambitions and in their employment of an aggressive behavior….Various traits of personality found their expression in differences in taste, mode of dress, hair style, use of cosmetics, the wearing of beard or of glasses" (italics in original).[192]

Juel-Nielsen's description differs dramatically from the numerous and widely reported anecdotal descriptions of MZA pairs provided by the MISTRA researchers, and by the popularizers of their work (more on this below).

Once again, missing data. The main personality inventories used in the MISTRA were the Minnesota Multiphasic Personality Inventory (MMPI), the California Psychological Inventory (CPI), the Multidimensional Personality Questionnaire (MPQ), and the Sixteen Personality Factor Questionnaire (16PF).

In a 1990 MISTRA publication of the CPI results, nine of the 20 CPI primary scale DZA group correlations were consistent with genetic predictions, seven did not differ significantly from (or were higher than) the corresponding MZA group correlation, and four DZA correlations were negative.[193] We do not know how the correlations fell in the MISTRA-administered 16PF personality study because they have never been published, even though the 16PF had been used extensively in personality research for decades. Segal did not report any 16PF results in her comprehensive *Born Together—Reared Apart*, even though she mentioned that the MISTRA "personality assessment forms…included the 16

Personality Factor Questionnaire."[194] As we have seen, the researchers denied others access to their raw data, and then published *selected* results that supported their pre-existing genetic biases, as well as the genetic biases and political agendas of their notorious funding sources.

"Personality" in dispute. Although it is widely known that IQ testing has been in dispute for over a century, the psychometric/behavioral genetic concept of a "personality trait" has also been controversial. Personality testing in behavioral genetic research is based on the assumption that "personality traits are relatively enduring individual differences in behavior that are stable across time and across situations," and can be measured and quantified with psychometric tests.[195] The psychometric/behavioral genetic emphasis on "individual differences" magnifies and emphasizes human behavioral differences, and tends to de-emphasize situational behavior and the common behaviors, abilities, longings and many other qualities that most human beings share. It attempts to "reduce our complicated, contradictory, changeable selves" to the "tidy label" of "personality," where it must be quantified for the purposes of genetic research.[196] Personality testing, as a pair of critics put it in 1990, "has never been uncontroversial in psychology."[197]

Segal acknowledged the critics' argument that personality inventories (tests) "do not faithfully capture the behaviors that people express in real life," and responded that although the inventories "are imperfect, they are superior to observational data gathered on just a few unrepresentative occasions."[198] It should be stressed, however, that "superior" does not equal "valid." Although it is not possible to "fake smart" on an IQ test, it is very possible to "fake good" or "fake bad" on a personality test, even though test developers usually create validity scales in an attempt to catch such faking. A group of organizational psychologists concluded in 2007 that "faking on self-report personality tests should be expected," that "it probably cannot be avoided," and that "corrections for faking do not appear to improve validity."[199]

Twins' answers on test questions. TRA researchers compare each twin's personality inventory score with a standardization (norm) group established by the test developers, after which twins' scores versus this standardization group are compared to each other. Researchers do not assess how many questions twins answer the same way. Theoretically, the members of a twin pair could answer individual questions very differently, yet their scale or total scores could be "highly correlated" based on similar raw or standardized scores. In a TRA personality study using an "empirically keyed" test such as the MMPI or the CPI, imagine an "Aggressiveness" scale based on 20 questions. Twin A might answer only the even-number questions (2, 4, 6… 20) in the keyed direction, while Twin B might answer only the odd-number questions (1, 3, 5…19) in the keyed direction, yet the pair would have correlated raw scores of 10 despite having answered each question differently. As one of the developers of a revised version of the MMPI recognized, "The same total raw score on a clinical scale can be achieved by individuals endorsing combinations of quite different kinds of items."[200]

Psychologist Paul Kline (1937-1999) noted in relation to empirically keyed self-report personality tests, "If two subjects have the same score on the scale, the scores are not necessarily psychologically equivalent."[201] In other words, TRA researchers interpret correlations on personality scales and inventories as showing that twins' personalities are "similar," even though they might not be so similar. In Kline's view the MMPI "is not a reliable or valid test," and "the validity of the CPI is dubious."[202] It would be interesting to see how similarly MZA and DZA pairs answered individual personality test questions, but this was not reported in the MISTRA publications.

Bouchard published a solo-authored 1994 article in *Science* on personality in the context of the MISTRA findings, where he made the remarkable claim that "the similarity we see in personality between biological relatives is almost entirely genetic in origin." Bouchard called this a "well-replicated finding in behavior genetics," whose "implications are straightforward."[203] He did not state what the "implications" of this dubious "finding" might be.

19. **The Results of a Little-Known Behavioral Genetic (Non-Twin) Adoption Study Contrasted Dramatically with the MISTRA Personality Findings.** In 1998, Robert Plomin and his Colorado Adoption Project (CAP) colleagues published the results of a longitudinal (non-twin) adoption study of personality. Plomin and colleagues found an average personality test-score correlation of .01 (that is, zero) between birthparents and their 240 adopted-away 16-year-old biological offspring, a correlation which they believed "directly indexes genetic influence."[204] Although they used their model-fitting analyses to calculate heritability and to claim genetic influences on personality traits (a textbook example of confirmation bias), the bottom line is that the results of this large and carefully planned adoption study showed no genetic influences on personality.[205]

In the 1990 MISTRA *Science* article, on the other hand, the researchers reported MZA personality correlations of .48-.50 based on 38-44 MZA pairs, and assumed/concluded that genetic factors caused these correlations. The 1998 CAP study, though flawed, was methodologically far superior to the MISTRA. Nevertheless, it is largely unknown. It is not often cited in academic publications (93 citations in Google Scholar as of October, 2018), and it is hardly ever mentioned in media reports or in books popularizing behavioral genetic research. As I showed in a 2013 analysis, in the few instances where the 1998 CAP study was cited in an academic publication, only a handful of authors informed their readers that Plomin et al. found no personality test-score correlation between birthparents and their adopted-away biological offspring. Plomin himself, though a prolific author and researcher, has not often cited his own 1998 CAP study.[206] In striking contrast, the 1990 MISTRA *Science* findings have been widely cited in academic publications for decades (1,885 citations in Google Scholar as of October, 2018), and several books have highlighted the study.

Segal noted that the results of the 1990 MISTRA *Science* article "appeared in hundreds of newspapers, magazines, and broadcasts across the country and around the world." And it is equally true that the results of the 1998 CAP study appeared in virtually no newspapers, in virtually no magazines, and in no broadcasts in the United States or anywhere else the

world. One could hardly imagine a better "natural experiment" showing widespread corporate media bias in favor of genetic explanations of human behavior.

20. Anecdotal Stories of Behaviorally Similar Twins Have "Nothing to Do with Genetics." Single-case anecdotal stories of allegedly very similarly behaving MISTRA MZA twins or triplets, such as the "Jim Twins," the "Fireman Twins," the "Nazi and Jew Pair," and the "Three Identical Strangers" provide no scientifically valid evidence in favor of genetic influences on behavior.[207] These and other sensationalized MISTRA-highlighted "cherry-picked" stories have been reported *ad nauseam* in the media since the late 1970s, and were featured in Segal's books. (It is noteworthy that Bouchard never published his own book about the study he headed, even as it was reported in a 1987 magazine article that he and Segal were working on one, due to be published in 1989.[208]) According to Wikipedia, cherry picking is "the act of pointing to individual cases or data that seem to confirm a particular position, while ignoring a significant portion of related cases or data that may contradict that position." As science writer and MISTRA critic John Horgan put it, "These tales of separated twins serve as a powerful rhetorical device, much more so than the statistical analyses and heritability figures."[209] Behaviorally *dis*similar MZA pairs are rarely mentioned in media reports or in textbooks.

Although these "spooky" and "eerie" stories have been highlighted in countless social and behavioral science textbooks over the past 30 years in the context of supporting genetic theories, the public has been massively misled by these selectively reported "bewitching science" tales.[210] As Richard Rose pointed out, the twins' similarities are likely the result of cohort effects, and have "nothing to do with genetics." Even the MISTRA researchers conceded the obvious point that "when any two biographies are avidly compared, at least some overlap is likely to be found."[211]

In 1984, the MISTRA-skeptical psychologist W. J. Wyatt and his colleagues assessed the similarities of 25 genetically *unrelated* pairs of college students matched on age and sex, and compared them with a group of MZTs. One pair of unrelated women had a lot in common: "Both are Baptist; volleyball and tennis are their favorite sports; their favorite subjects in school were English and math (and both listed shorthand as their least favorite); both are studying nursing; and both prefer vacations at historical places."[212] As Juel-Nielsen pointed out, twin researchers and others are "consciously or unconsciously, in a position to choose to emphasize similarities between the twins, and at the same to omit to register, or be inclined to belittle, the differences."[213]

We saw earlier that in the 2018 movie *Three Identical Strangers*, one of the triplets admitted that he and his reunited brothers had emphasized their similarities, and this certainly occurred in other celebrated cases. The triplets' story is entertaining, fascinating, anger provoking, and tragic. Their obvious motive in the 1980s was fame and fortune, and they cannot be faulted for attempting to seize on their opportunity. But influential authors claiming that their story proves something about genetic influences on behavior *can* be faulted. Sadly, the triplets were separated at six months of age as part of Neubauer's harmful and misguided "secret study," and later became the unwitting pawns of "scientists"

and others attempting to use their story to sell dubious genetic determinist theories of behavior to the similarly unwitting general public.

Insofar as the MISTRA-supplied or promoted stories of reunited MZ pairs or triplets are put forward as scientific evidence in favor of the claim that genetic factors influence human behavioral similarities, they belong in the pseudoscience category currently inhabited by Bigfoot tracking, ghost sightings, astrology, psychic surgery, and other such areas. These selectively reported stories are appropriate mainly for TV celebrity shows, and for the tabloids and magazines we pass by in the supermarket checkout lane.

21. **There is a Lack of Accountability and Pre-Registration in Social and Behavioral Science Research.** Problematic research practices and reporting occurred not only in the MISTRA, but occur frequently in psychology and in the social and behavioral sciences in general. In *The Seven Deadly Sins of Psychology*, Chambers pointed to many problem areas in the research/publication process in psychology and other fields.[214] He provided a valuable framework for describing biased, deceptive, and even fraudulent research practices.

    In 1981, Farber called for the creation of a central registry for TRA data that would act as a safeguard "against the disreputable claims and use of data that have occurred in the past and undoubtedly will occur in the future."[215] Unfortunately, the TRA data registry that Farber envisioned was never created. Fortunately, a movement is now underway to make pre-registration the norm in the social and behavioral sciences. Although "we may never be able to eliminate bias altogether from human nature," Chambers wrote, a "sure way to immunize ourselves against its consequences…is peer-reviewed study preregistration."[216]

    Like most other social and behavioral science research projects, there exists no pre-registered publically accessible account of the MISTRA researchers' practices, or a description of how they planned to interpret their findings at various stages. This allowed them to present their studies in neatly packaged publications, with little prior record of their intended methods, assumptions, definitions, comparisons, and decision-making processes.

22. **"Genes for Behavior" Are Still "Missing."** The year 2018 marked the dubious 10th anniversary of the "missing heritability" era. The claim that genes for behavior and psychiatric disorders are "missing," as opposed to concluding that these genes do not exist, is an attempt to explain the failure to make confirmed discoveries of genetic variants that cause differences in IQ scores, personality, and other forms of behavior by claiming that such variants exist and await discovery once better methods are found, and larger samples are obtained. Based on the common claim that the MISTRA and other behavioral genetic studies have confirmed the "moderate to high heritability" of general intelligence (*g*) and personality, molecular genetic researchers, armed with generous funding and modern gene-finding technology, should be able to identify the genes that underlie it. Bouchard admitted in 2014, however, that although "the high heritability of *g* [general intelligence] has made it a popular target in the search for genes that influence behavior," the "results to date have been dismal in comparison with expectation."[217] We have seen many gene discovery claims for general intelligence, personality, psychiatric

disorders, and other behavioral characteristics since the 1990s, yet follow-up studies were unable to confirm the original findings.

In a 2018 article, Plomin claimed that genes for intelligence had been identified the previous year. At the same time, he acknowledged that attempts up to 2017 had failed: "Similar to results for many other complex traits, early results for intelligence were disappointing for more than 100 candidate gene studies and for seven GWAS [genome-wide association studies]. From the 1990s until 2017, no replicable associations were found."[218] History has shown that it is extremely unlikely that these recent claims will be confirmed, and Plomin has a 40-year track record of making claims of behavioral gene discoveries that turned out to be false-positive findings that could not be replicated.[219]

Given the decades of sensationalized yet non-replicated claims of gene discoveries for behavioral traits and psychiatric disorders, the media response to recent claims based on genome-wide association studies, or to newer studies based on the "polygenic risk score" method, should be extreme skepticism and caution similar to the "oh no, not again" skepticism and caution *Peanuts* comic strip character Charlie Brown responded with whenever Lucy van Pelt asked him to kick the football she was holding. Until proven otherwise, we should assume that current and future behavioral gene discovery claims are false-positive findings. As Ken Richardson described it in his 2017 book *Genes, Brains, and Human Potential: The Science and Ideology of Intelligence*, the IQ "gene bubble" is "bursting."[220]

In a 1998 article, McGue and Bouchard commented on what was for them, but not for many critics, the unexpected failure to identify genes that cause behavioral differences:

> "The failure to identify the genes underlying specific human behavioral phenotypes may indicate that we have been misled by the twin and adoption study findings. Alternatively, and we believe more plausibly, the current failure may simply reflect the difficulty of gene identification with complex and heterogeneous phenotypes. Additional molecular genetic research should provide the necessary observations to resolve these two possibilities."[221]

Two decades later, the failure of "additional molecular genetic research" to discover "genes underlying specific human behavioral phenotypes" has helped "resolve" the issue in favor of the first possibility: *We have indeed been misled by twin studies, by adoption studies, and by twin-adoption (TRA) studies such as the MISTRA.*

Conclusions

I have listed and discussed 22 reasons why we must reject the Minnesota Study of Twins Reared Apart researchers' conclusion that their study found genetic influences, major or otherwise, on IQ, "special mental abilities," "personality," and other forms of behavior. The major problem areas I discussed include:

- The MISTRA recruitment methods favored the inclusion of behaviorally more similar twin pairs
- It is likely that most MZA pairs were only partially reared apart
- The generalization of the results to the non-twin population is questionable
- Even perfectly separated MZA pairs experience many common non-familial environmental influences, including cohort influences, which the MISTRA researchers assumed did not exist
- The findings were based on accepting the validity of controversial psychometric/behavioral genetic concepts such as "IQ," "general intelligence," "personality," "model fitting," and "heritability"
- The researchers completely removed the DZA control group results and correlations from their 1990 *Science* IQ study, most likely because the full-sample MZA and DZA group IQ correlations did not differ at statistically significant levels
- The MISTRA full-sample DZA control group IQ correlations remain unpublished to this day
- The researchers bypassed two steps in their data analysis procedure (Step 2 and Step 3A) in the process of determining whether genetic factors influence IQ scores
- The researchers' model-fitting analyses were based on a string of false or unsupported assumptions, some of which they admitted "are likely not to hold"
- The researchers mistakenly counted environmental influences on MZA behavioral resemblance as genetic influences, thereby creating a genetic "heads I win, tails you lose" type of study that guaranteed that genetic interpretations of above-zero MZA group psychological test-score correlations would prevail
- There was confirmation bias in the study, which flowed from the researchers' pre-existing beliefs about the importance of genetic influences on intelligence, and on human behavioral differences in general
- A software program used by the researchers was biased in favor of producing genetic findings
- The researchers accepted the twins' potentially unreliable accounts of their degrees of separation, contact, and behavioral similarity
- The researchers failed to provide adequate life history and test-score information for the twins
- The researchers failed to provide access to the MISTRA raw data
- Highly publicized anecdotal stories of individual MZA pairs provide no evidence in support of genetic influences on behavior

- We are witnessing the ongoing decades-old failure to make confirmed discoveries of genetic variants that cause differences in IQ scores, personality, and other forms of behavior

The central fallacy of the MISTRA was that the researchers claimed that above-zero MZA group psychological test-score correlations were caused by shared genetic influences, when these correlations can be explained by cohort effects and other *non*-genetic influences and biases, plus researcher bias and error. Like the TRA and twin method studies that came before it, the MISTRA findings were confounded by environmental influences, and the researchers were therefore unable to disentangle the potential influences of genes and environments on human behavior. This leads to the conclusion that the study's findings must be rejected, that textbooks should be rewritten to reflect this conclusion, and that this should be widely reported in the media.

The researchers claimed that any genetic biases they may have had did not influence how they analyzed the data, or how they arrived at their conclusions. According to Bouchard and McGue, "Behavior genetic methods are unbiased with regard to whether genetic or environmental sources of variance are more important," and that "if there is no genetic source of variance the methods will reveal this fact."[222] And according to Segal, "We were interested in results of any kind on any topic that was studied. We did not decide how the data turned out, *the twins did*" (italics in original).[223] However, the "twins" did not decide to base the study's findings on key assumptions that are "likely not to hold," to omit the DZA control group results and correlations from the IQ study, to count environmental influences as genetic influences, to suppress the full-sample DZA group IQ correlations, to find a genetic explanation for DZA correlations that did not fit genetic predictions, to use a genetically biased computer software program to produce genetic findings, to hope that a number of false assumptions "offset each other" in favor of genetics, and to deny critics access to the raw data—*the researchers did*. Their strong genetic biases, it seems, compelled them to omit, bypass, and suppress their control group DZA IQ correlations in order to obtain the desired results.

The MISTRA samples, methods, and findings have been systematically misinterpreted, misrepresented, and misreported in psychology textbooks and other influential secondary sources. Psychology textbook authors often endorse the MISTRA researchers' conclusions with little critical analysis, and frequently reproduce the highly misleading anecdotal stories and photographs of pairs selectively released by the MISTRA researchers. Adding to this the American corporate media's scandalously uncritical celebration of the MISTRA since the early 1980s, and the many popular books about behavioral genetics and twin research that have appeared during the past thirty years, we are witnessing the latest Internet-era version of the "twins reared apart" deception that Leon Kamin exposed in the 1970s. I have attempted to update and further expose this deception.

In a healthy and vibrant academic field, as opposed to a stagnant and regressive one, a combination of the suppression of the full-sample DZA control group IQ correlations and the omission of those correlations in the IQ study, unethical "data hoarding" practices, the great social and political importance of the study, the genetic determinist claims made on the basis of it (including those by eugenics and white nationalist groups, which includes the main MISTRA

funding source), and the overturning of previous research—and plain common sense—showing the power of environmental influences, might have led to the creation of a commission to objectively review the study in great detail in order to confirm or disconfirm the findings. Instead, the American Psychological Foundation presented Bouchard with its 2014 "Gold Medal Award for Lifetime Achievement in the Science of Psychology," which cited the MISTRA as "groundbreaking and inventive, exciting and controversial," and a "stunning achievement, a body of work in which all psychologists can take pride."[224] Bouchard also received the 2018 "Dunnette Prize for the Study of Individual Differences" from The Society for Industrial and Organizational Psychology, which is affiliated with the U.S. Association for Psychological Science. Segal's *Born Together—Reared Apart* won the American Psychological Association's 2013 "William James Book Award." As three authors discussing recent "underperforming big ideas in research" pointed out, "Criteria such as number of publications, citations, prizes, and recognition are irrelevant as these are simply self-rewarding artifacts of the system."[225] The "system" in American academic psychology and other social and behavioral science fields, at least as it relates to the evaluation of twin studies and other areas of behavioral genetic research, is clearly broken.

As stated earlier, when social and behavioral science investigators refuse to make their raw data available for inspection and analysis by qualified reviewers—especially when the study's results have important social policy implications—we must automatically reject their findings when they are based on samples taken from rare populations that cannot be independently reproduced, such as reared-apart twins. The MISTRA researchers did not allow independent review of their raw data, and then asked us to accept their claim that genetic factors play a major role in causing behavioral differences among billions of human beings—past, present, and future— on the basis of how a few hundred abandoned twins performed on psychological tests. We must decisively reject this claim.

About the Author

Jay Joseph, Psy.D. is a clinical psychologist practicing in the San Francisco Bay Area. Since 1998, he has published peer-reviewed articles and book chapters critically examining genetic theories and research in the social and behavioral sciences. He is the author of *The Gene Illusion: Genetic Research in Psychiatry and Psychology Under the Microscope* (2004), *The Missing Gene: Psychiatry, Heredity, and the Fruitless Search for Genes* (2006), *The Trouble with Twin Studies: A Reassessment of Twin Research in the Social and Behavioral Sciences* (2015), and *Schizophrenia and Genetics: The End of an Illusion* (2017). His blog "The Gene Illusion" can be found at the Mad in America website. A complete list of his publications can be found at his website.

Acknowledgements

I thank M.C. Jones and Ken Richardson for providing helpful feedback on earlier versions of this analysis. At the same time, all opinions expressed and conclusions reached in this article are my own, and do not necessarily reflect the views of the people who assisted me. Any possible errors are entirely my responsibility.

ENDNOTES

[1] Chambers, C., (2017), *The Seven Deadly Sins of Psychology: A Manifesto for Reforming the Culture of Scientific Practice*, Princeton, NJ: Princeton University Press, p. 175.

[2] Hogben, L., (1933), *Nature and Nurture,* London: George Allen & Unwin, p. 121.

[3] Bouchard, T. J., Jr., (1998), Genetic and Environmental Influences on Adult Intelligence and Special Mental Abilities, *Human Biology*, *70*, 257-279, pp. 259-260.

[4] Wright, L., (1997), *Twins: And What They Tell Us About Who We Are*, New York: John Wiley & Sons, pp. 143-144.

[5] See Joseph, J., (2004), *The Gene Illusion: Genetic Research in Psychiatry and Psychology Under the Microscope*, New York: Algora; Joseph, J., (2006), *The Missing Gene: Psychiatry, Heredity, and the Fruitless Search for Genes*, New York: Algora.

[6] Joseph, J., (2015), *The Trouble with Twin Studies: A Reassessment of Twin Research in the Social and Behavioral Sciences*, New York: Routledge.

[7] Segal, N. L., (2012), *Born Together—Reared Apart: The Landmark Minnesota Twin Study*, Cambridge, MA: Harvard University Press, Appendix A.

[8] Tucker, W. H., (2013), [Review of the Book *Born Together—Reared Apart: The Landmark Minnesota Twin Study*, by N. L. Segal], *Journal of the History of the Behavioral Sciences*, *49*, 337–341, p. 339.

[9] Segal, 2012, p, 317.

[10] Segal, 2012, Appendix A.

[11] Tucker, 2013, p. 337.

[12] Segal, 2012, p. 400.

[13] Research performed in the 21st century has called into question the basic twin study assumption that both members of an MZT or MZA pair are genetically identical throughout their lives; see Charney, E., (2012), Behavior Genetics and Postgenomics, *Behavioral and Brain Sciences*, *35*, 331-358.

[14] Segal, 2012, p. 42. The MISTRA publications most often did not report separate correlations for same- and opposite-sex DZAs.

[15] Plomin et al., (2013), *Behavioral Genetics* (6th ed.), New York: Worth Publishers, p. 87.

[16] Plomin et al., 2013, pp. 86-87.

[17] Joseph, 2015, Chapter 7. See also Joseph, J., (1998), The Equal Environment Assumption of the Classical Twin Method: A Critical Analysis, *Journal of Mind and Behavior, 19*, 325-358; Joseph, J., (2013a), The Use of the Classical Twin Method in the Social and Behavioral Sciences: The Fallacy Continues, *Journal of Mind and Behavior*, *34*, 1-39.

[18] Bouchard, T. J., Jr., Lykken, D. T., McGue, M., Segal, N., & Tellegen, A., (1990), Sources of Human Psychological Differences: The Minnesota Study of Twins Reared Apart, *Science, 250,* 223-228. (From here on, referenced as "Bouchard, Lykken, et al., 1990.")

[19] Quoted in Segal, 2012, p. 104.

[20] Segal, 2012, p. 112.

[21] Bouchard, Lykken, et al., 1990, p. 223.

[22] Bouchard, Lykken, et al., 1990, p. 224.

[23] Segal, 2012, p. 105.

[24] Segal, 2012, p. 104.

[25] McGue et al., (1993), "Behavioral Genetics of Cognitive Ability: A Life-Span Perspective," in R. Plomin & G. McClearn (Eds.), *Nature, Nurture, and Psychology* (pp. 59-76), Washington, DC: American Psychological Association; Bouchard, T. J., Jr., (1993b), "Genetic and Environmental Influences on Adult Personality: Evaluating the Evidence," in J. Hettema & I. Deary (Eds.), *Basic Issues in Personality* (pp. 15-44), Dordrecht, The Netherlands: Kluwer Academic Publishers.

[26] Bouchard, Lykken, et al., 1990, pp. 223, 227.

[27] Bouchard, T. J., Jr., (1997), "IQ Similarity in Twins Reared Apart: Findings and Responses to Critics," in R. Sternberg & E. Grigorenko (Eds.), *Intelligence, Heredity, and Environment* (pp. 126-160), New York: Cambridge University Press, p. 130.

[28] Juel-Nielsen, N., (1980), *Individual and Environment: Monozygotic Twins Reared Apart* (revised ed.), New York: International Universities Press (originally published in 1965); Newman et al., (1937), *Twins: A Study of Heredity and Environment,* Chicago, University of Chicago Press; Shields, J., (1962), *Monozygotic Twins Brought Up Apart and Brought Up Together,* London: Oxford University Press.

[29] Plomin et al., 2013, p. 412.

[30] Segal, 2012, p. 63.

[31] Segal, 2012, p. 343, Note 69.

[32] Segal, 2012, p. 12.

[33] Bouchard et al., (1986), "Development in Twins Reared Apart: A Test of the Chronogenetic Hypothesis," in A. Demirjian (Ed.), *Human Growth: A Multidisciplinary Review* (pp. 299-310), London: Taylor & Francis, p. 300. Although at the beginning of the study Bouchard had designated DZA pairs as the MISTRA control group, and even though he had reaffirmed this in this 1986 publication, in his 2018 "Dunnette Prize Address" he implied (video starting at about 50:38) that the MISTRA control group consisted of *MZT* pairs: "Once we got our study going we realized…that we needed comparison groups, ordinary twins, non-adoptees and so we built a huge twin registry….We immediately built a control comparison group to see if we could replicate our findings." (Retrieved from YouTube 10/23/2018, URL: https://www.youtube.com/watch?v=zAPBHis9GKg).

[34] McGue, M., & Bouchard, T. J., Jr., (1989), "Genetic and Environmental Determinants of Information Processing and Special Mental Abilities: A Twin Analysis," in R. Sternberg (Ed.), *Advances in the Psychology of Human Intelligence* (Vol. 5, pp. 7-45), Hillsdale, NJ: Erlbaum, p. 22.

[35] Tellegen et al., (1988), Personality Similarity in Twins Reared Apart and Together, *Journal of Personality and Social Psychology, 54,* 1031-1039, 1988, p. 1034.

[36] Bouchard, T. J., Jr., (1984), "Twins Reared Together and Apart: What They Tell Us about Human Diversity," in S. Fox (Ed.), *Individuality and Determinism: Chemical and Biological Bases* (pp. 147-184), New York: Plenum Press, p. 150.

[37] In his discussion of the 1990 MISTRA *Science* study, the author of *Forty Studies That Changed Psychology* made the error of writing that the MISTRA genetic conclusions were based on MZA-MZT correlations, as did Hamer in *Living with Our Genes*. Hock, R. R., (2009), *Forty Studies That Changed Psychology: Explorations into the History of Psychological Research* (6th ed.), Pearson: Upper Saddle River, NJ.

[38] Although in their research publications Bouchard and colleagues did not base their genetic conclusions on MZA versus MZT comparisons, there are a few instances in later commentaries where they stated or implied that similar MZA-MZT correlations support genetic conclusions. As one example, see Lykken, 1995

[39] Segal, N. L., (2017), *Twin Mythconceptions: False Beliefs, Fables, and Facts About Twins,* London: Academic Press, p. 32.

[40] Bouchard, Lykken, et al., 1990, p. 227. In 1982 Lykken described a hypothetical scenario in which the newborn babies of U.S. 1960s presidential candidates Barry Goldwater (conservative) and Hubert Humphrey (liberal) had been switched at birth. If their "sons had been mixed up in the hospital nursery," Lykken wrote, "who could doubt that their political sentiments would have followed those of their adoptive parents?" Doesn't Lykken's thought experiment suggest that the influence of the "shared family environment" is important? See Lykken, D. T., (1982), Research with Twins: The Concept of Emergenesis, *Psychophysiology, 19,* 361-373, p. 372.

[41] Segal, 2012, p. 62.

[42] Segal, 2012, p. 111.

[43] Segal, 2012, p. 61.

[44] Bouchard, Lykken, et al., 1990, p. 223.

[45] Bouchard, 1997, p. 129.

[46] John et al., (2012), Measuring the Prevalence of Questionable Research Practices with Incentives for Truth Telling, *Psychological Science, 23,* 524-532.

[47] Chambers, 2017, p. 25.

[48] Walster, G. W., & Cleary, T. A., (1970), A Proposal for a New Editorial Policy in the Social Sciences, *The American Statistician, 24,* 16-19, p. 18.

[49] Kamin, L. J., (1974), *The Science and Politics of I.Q.*, Potomac, MD: Erlbaum.

[50] Pedersen et al., (1988), Neuroticism, Extraversion, and Related Traits in Adult Twins Reared Apart and Reared Together, *Journal of Personality and Social Psychology, 55,* 950-957, p. 955.

[51] Newman et al., 1937, p. 31.

[52] Segal, 2012, pp. 35-36.

[53] Shields, 1962, p. 6.

[54] Lykken et al., (1978), Volunteer Bias in Twin Research: The Rule of Two-Thirds, *Social Biology, 25,* 1-9, p. 6.

[55] Rosenthal, R., & Rosnow, R. L., (1975), *The Volunteer Subject*, New York: John Wiley & Sons.

[56] Segal, 2012, p. 42; see also Lykken et al., 1978.

[57] Farber, S. L., (1981), *Identical Twins Reared Apart: A Reanalysis*, New York: Basic Books, p. 36.

[58] Farber, 1981, p. 8.

[59] Bouchard et al., 1986, p. 301.

[60] Lush, J. L., (1949), Heritability of Quantitative Characteristics in Farm Animals, *Hereditas* (Suppl.), G. Bonnier & R. Larsson (Eds.).

[61] Criticism of the "heritability" concept can be found in many places, including Chaufan, C., (2008), Unpacking the Heritability of Diabetes: The Problem of Attempting to Quantify the Relative Contributions of Nature and Nurture, *DataCrítica: International Journal of Critical Statistics, 2,* 23-38; Keller, E. F., (2010), *The Mirage of a Space Between Nature and Nurture*, Durham, NC: Duke University Press; Hirsch, J., (1997), Some History of Heredity-vs-Environment, Genetic Inferiority at Harvard (?), and The (Incredible) Bell Curve, *Genetica, 99,* 207-224; and Stoltenberg, S. F., (1997), Coming to Terms with Heritability, *Genetica, 99,* 89-96.

[62] Moore, D. S., & Shenk, D., (2016), The Heritability Fallacy, *WIREs Cognitive Science*, 1-8, DOI: 10.1002/wcs.1400

[63] Wahlsten, D., (1990), Insensitivity of the Analysis of Variance to Heredity-Environment Interaction, *Behavioral and Brain Sciences, 13,* 109-120.

[64] Johnson et al., (2009), Beyond Heritability: Twin Studies in Behavioral Research, *Current Directions in Psychological Science, 18,* 217-220, p. 220.

[65] Leon J. Kamin, personal communication of 4/8/2001. For more on the importance of sharing data, and on the MISTRA researchers' practice of denying access to their raw data to Kamin and other critics, see Joseph, 2015, pp. 122-126.

[66] Wright, 1997, p. 69.

[67] Rose, S., (1997), *Lifelines: Life Beyond the Genes*, New York: Oxford University Press, p. 207; Tucker, W. H., (1997), Re-Reconsidering Burt: Beyond a Reasonable Doubt, *Journal of the History of the Behavioral Sciences, 33,* 145-162, p. 160, footnote 65.

[68] A description of "data hoarding" practices in psychology and other fields can be found in Chambers, 2017. The American Psychological Association's 2002 "Ethical Principles of Psychologists and Code of Conduct" can be found at https://memforms.apa.org/apa/cli/interest/ethics1.cfm#8_14

[69] Levelt Committee, Noort Committee, Drenth Committee, (2012), *Flawed Science: The Fraudulent Research Practices of Social Psychologist Diederik Stapel*, p. 58.

[70] Segal, 2012, p. 184.

[71] Segal, 2012, p. 299.

[72] Turning the MISTRA story on its head, imagine a hypothetical team of environmentally oriented researchers from the University of Nebraska, who perform a TRA study called the "Nebraska Urban Research on Twins Using Regulation of Emotions Study" ("NURTURES"). Based on their twin data, the researchers publish a widely reported paper in *Science* showing conclusively that environmental factors have a "pronounced and pervasive influence on behavioral variability," and that genetic factors play only a minor role. The NURTURES researchers conclude that their results apply (generalize) to the general population, and then call for a radical reevaluation of the causes of human behavior in favor of the environment. They guard their raw data and information closely, they deny access to it to genetically oriented critics, and they publish only information and results that they choose to publish. Textbooks, popular writers, and the media cover this study extensively, with most accounts claiming that the NURTURES study showed conclusively that environmental (non-genetic) factors are the main cause of differences in human behavior and intelligence. This position would be supported by non-TRA research, and by the widespread media reporting of anecdotal stories of individual MZA pairs, selectively released by the NURTURES researchers, showing the "eerie" fact that these genetically identical pairs had almost nothing in common. Most likely, the NURTURES study would be severely criticized by genetic researchers using some of the arguments currently put forward by critics of TRA studies.

[73] Holden, C., (2009), Behavioral Geneticist Celebrates Twins, Scorns PC Science, *Science, 325,* 27; Anonymous, (2014), Gold Medal Award for Life Achievement in the Science of Psychology: Thomas J. Bouchard, Jr., Award for Distinguished Scientific Contributions, *American Psychologist, 69,* 477-479, p. 477.

[74] Anonymous, (1974, April 4th), Freedom of Expression (editorial), *Minnesota Daily,* p. 4.

[75] Bouchard, 1976, p. 193.

[76] Bouchard, T. J., Jr., (1995), Breaking the Last Taboo [Review of the Book *The Bell Curve: Intelligence and Class Structure in American Life*, by R. J. Herrnstein & C. Murray], *Contemporary Psychology, 40, (5),* 415-421.

[77] Psychologist J. Philippe Rushton was the Director of the Pioneer Fund from 2002 until his death in 2012. Rushton was well known for his belief, which he claimed was backed by scientific research, that people of African descent are genetically inferior to Caucasians in measured IQ and other characteristics (and that East-Asians are superior to both). Rushton first published his book *Race, Evolution, and Behavior: A Life History Perspective* in 1995, which was followed by several revised editions. In a 2000 abridged edition of *Race, Evolution, and Behavior*, Rushton recounted the "shocked" observations of early European "explorers" in Africa (actually, in most cases colonial plunderers and conquerors attempting to justify murder, plunder, and conquest), while approvingly listing some of the crudest racial stereotypes imaginable, such as comparing African's behavior to "wild animals," their having a "natural sense of rhythm," and their being prone to cannibalism. Rushton was implying that people of the African diaspora currently continue to harbor similar allegedly inherited tendencies. See Rushton, J. P., (2000), *Race, Evolution, and Behavior: A Life History Perspective* (2nd Special Abridged Ed.), Port Huron, MI: Charles Darwin Research Institute, p. 8. In recommending Rushton's book, Bouchard wrote, "The only acceptable explanation of race differences in behavior allowed in public discourse is an entirely environmental one...Professor Rushton deserves our gratitude for having the courage to declare that 'this emperor has no clothes,' and that a more satisfactory explanation must be sought." See Bouchard, T. J., Jr., in Rushton, 2000, p. 2.

[78] Lombardo, P. A., (2002), "The American Breed": Nazi Eugenics and the Origins of the Pioneer Fund, *Albany Law Review, 65,* 743-830; Tucker, W. H., (2002), A Closer Look at the Pioneer Fund: Response to Rushton, *Albany Law Review, 66,* 1145-1159.

[79] Lykken, D. T., (2000), The Causes and Costs of Crime and a Controversial Cure, *Journal of Personality, 68,* 559-605.

[80] Lykken, D. T., (1995), *The Antisocial Personalities*, Hillsdale, NJ: Erlbaum, p. 233.

[81] Lykken, D. T., (2004), The New Eugenics [Review of the Book *Eugenics: A Reassessment*, by R. Lynn], *Contemporary Psychology, 49,* 670-672, p. 672. In his 2001 book *Eugenics: A Reassessment* (published by Praeger), which Lykken praised in his review, psychologist Richard Lynn provided a detailed argument in support of instituting eugenic breeding plans to improve "health, intelligence, and moral character," and called for the restoration of compulsory eugenic sterilization and other negative eugenic measures (Lynn, 2001, p. 53). Lynn supported the compulsorily eugenic sterilization of "criminals" and others, writing that "eugenics does not require the extermination of undesirables. It is sufficient for eugenics that the mentally retarded and recidivist criminals should be sterilized" (p. 239). Lynn discussed the possibility of injecting a virus into the bodies of low IQ scorers and "psychopaths" at age 12 as a means of "temporary sterilization," in the same way this method has been used for "the sterilization of rabbits, kangaroos, and other pests" (p. 213). At age 22, this person could apply for parental license. If granted, the sterilization would be reversed. If not, the person "could be vasectomized" (p. 213). "In time," wrote Lynn, "the mandatory sterilization of boys might come to seem no more objectionable than inoculations against infectious diseases" (p. 212). He believed that the "elimination" of the "underclass" found in Western societies "must be one of the objectives of eugenics" (p. 133), and that the "sterilization of the socially undesirable" could be justified on the same grounds as the "prohibition of smoking in public places and the discharge of poisons into rivers and of pollutants into the atmosphere" (p. 241). Lynn regarded Adolf Hitler as someone whose "objective was to improve the genetic quality of the Germans" (p. 54), thereby accepting the Nazi definition of "Germans" as meaning "Aryan" Germans only, since their aim was to exterminate, enslave, or dominate all non-"Aryans." Throughout his book it is clear that Hitler's Germany was for Lynn an example, though an imperfect one, of how the eugenic program could be achieved. He made neutral or positive references to actual German National Socialist politician and scientist war criminals who played a major role in carrying out the Holocaust, such as Hitler, racial hygienist twin researcher Otmar von Verschuer (pp. 28, 142), and Heinrich Himmler (p. 216). Lynn believed that it would be difficult to establish a eugenic society in a "democratic state," and clearly preferred the establishment of "authoritarian eugenic states" (p. 299) in the West in order to suppress those groups that surely would oppose the implementation of eugenic measures. Among the groups that "would be expected to oppose any attempt to rehabilitate eugenics," according to Lynn, are the "racial and ethnic minorities that would be disproportionality affected by eugenic policies…such as African-American and Hispanics in the United States whose low average intelligence and high crime rates would make them disproportionately subject to sterilization and restrictions on immigration" (p. 277). In his review, Lykken called Lynn's *Eugenics: A Reassessment* an "excellent, scholarly book" (Lykken, 2004, p. 672).

[82] Lykken, 1995, pp. 216-217.

[83] Lykken, 1995, p. 71.

[84] Rossi, M., (2013), Consider Jack and Oskar, [Review of the book *Born Together—Reared Apart: The Landmark Minnesota Twin Study*, by N. L. Segal], *London Review of Books, 35 (3),* 346-349.

[85] Bouchard, T. J., Jr., (Interviewee), (2004), Distinguished Contributor Interview: Thomas Bouchard, Jr., International Society for Intelligence Research (ISIR) Conference 2004 (Part 1), retrieved from the ISIR Web site on 7/28/2013 from http://www.isironline.org/interviews/interview2004.html

[86] Segal, 2012, p. 305.

[87] Mukherjee, S., (2016), *The Gene: An Intimate History*, New York, Scribner, p. 383.

[88] My three 2016 articles documenting Mukherjee's inaccurate descriptions in *The Gene: An Intimate History* are as follows: Joseph, J., (2016, June 2), Reared-Apart Twin Study Mythology: The Latest Contribution (Part One), [Web log post, *Mad in America* "The Gene Illusion"]; Joseph, J., (2016, June 27), Reared-Apart Twin Study Mythology: The Latest Contribution (Part Two), [Web log post, *Mad in America* "The Gene Illusion"]; Joseph, J., (2016, July 12), Schizophrenia Twin Research as Reported in *The Gene: An Intimate History*: Getting the Facts Straight, [Web log post, *Mad in America* "The Gene Illusion"].

[89] Segal, 2012, p. 15.

[90] Beckwith et al., (1991), IQ and Heredity [Letter to the editor], *Science, 252,* 191.

[91] Bouchard et al., (1991), Response [Letter to the editor], *Science, 252,* 191-192.

[92] In *Born Together—Reared Apart*, Segal (2012, pp. 113-114) reviewed the issues raised by Beckwith and others, and claimed that the "University of Minnesota's institutional review board (like all university review panels) requires confidentiality of participant data. This was stated in our protocol, and the twins signed a consent form acknowledging their understanding of this agreement." However, there are many instances in which the MISTRA researchers published or provided journalists with "participant data," and many stories about various pairs are found in Segal's books. In fact, her book *Indivisible by Two* consisted mainly of (selectively reported) stories about twins. In most cases, the information she provided supported the researchers' genetic conclusions.

[93] Bouchard, T. J., Jr. & Juel-Nielsen, N., (1981), Interview (Video), posted on YouTube 8-5-2018, retrieved 8-19-2018 from https://www.youtube.com/watch?v=aqwXFrV7-Z0

[94] Farber, 1981, p. 62.

[95] Cassou et al., (1980), Génétique et Schizophrénie: Réévaluation d'un Consensus [Genetics and Schizophrenia: Reevaluation of a Consensus], *Psychiatrie de l'Enfant, 23,* 87-201.

[96] Shore, A. N., (2001), Effects of a Secure Attachment Relationship on Right Brain Development, Affect Regulation, and Infant Mental Health, *Infant Mental Health Journal, 22,* 7-66; Siegel, D. J., (1999), *The Developing Mind: Toward a Neurobiology of Interpersonal Experience*, New York: Guilford.

[97] Newman et al., (2015), Attachment and Early Brain Development—Neuroprotective Interventions in Infant-Caregiver Therapy, *Translational Developmental Psychiatry, 3*: DOI: 28647- http://dx.doi.org/10.3402/tdp.v3.28647

[98] Langinvainio et al., 1981, "Finnish Twins Reared Apart: Preliminary Characterization of Rearing Environment," in L. Gedda et al., (Eds.), *Twin Research 3: Part B. Intelligence, Personality, and Development* (pp. 189-198), New York: Alan R. Liss, p. 198.

[99] Bouchard, T. J., Jr., (1976), "Genetic Factors in Intelligence," in A. Kaplan (Ed.), *Human Behavior Genetics* (pp. 164-197), Springfield, IL: Charles C. Thomas, p. 173.

[100] Bouchard, T. J., Jr., (1993a), "The Genetic Architecture of Human Intelligence," in P. Vernon (Ed.), *Biological Approaches to the Study of Human Intelligence* (pp. 33-93), Norwood, NJ: Ablex Publishing Corporation, p. 56.

[101] Fancher, R. E., (1985), *The Intelligence Men: Makers of the IQ Controversy*, New York: W.W. Norton, p. 165.

[102] Sternberg, R. J., (2007), "Critical Thinking in Psychology is Really Critical," in R. Sternberg et al., (Eds.), *Critical Thinking in Psychology* (pp. 289-296), New York: Cambridge University Press, p. 292.

[103] Bouchard, T. J., Jr., (1987), "The Hereditarian Research Program: Triumphs and Tribulations," in S. Modgil & C. Modgil (Eds.), *Arthur Jensen: Consensus and Controversy* (pp. 55-75), New York: Falmer Press, p. 65; Bouchard, 1998, p. 261.

[104] Stoolmiller, M., (1999), Implications of the Restricted Range of Family Environments for Estimates of Heritability and Nonshared Environment in Behavior-Genetic Adoption Studies, *Psychological Bulletin, 125,* 392-409, p. 404.

[105] For a description of the environmental similarities, levels of contact, and mutual relationship of all 75 MZA pairs reported in the Newman, Shields, and Juel-Nielsen studies, see Joseph, 2015, Chapter 2, Tables 2.1, 2.2, & 2.3. See also Kamin, 1974; Taylor, 1980.

[106] The other two modern TRA studies are a study from Finland, and the Swedish Adoption/Twin Study on Aging (SATSA). In the Swedish study the investigators considered pairs to have been "reared-apart" if they had been separated before age 11: "By definition, the twins reared apart were separated by the age of 11" (Pedersen et al., 1992, A Quantitative Genetic Analysis of Cognitive Abilities During the Second Half of the Life Span, *Psychological Science, 3,* 346-353, p. 347). About 75% of the SATSA pairs had some degree of contact after separation. According to the researchers, the SATSA twins (average age 65.6 years) were "separated" for an average of only 10.9 years at the time of testing (Pedersen et al., 1992, p. 347). 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 5, and the degree of post-separation contact is unclear (Langinvainio et al., 1984, Finnish Twins Reared Apart II: Validation of Zygosity, Environmental Dissimilarity and Weight and Height, *Acta Geneticae Medicae et Gemellologiae, 33,* 251-258).

[107] Segal, 2012, p. 37.

[108] Segal, 2012, p. 38.

[109] Sapolsky, R. M., (2017), *Behave: The Biology of Humans at Our Best and Worst*, New York: Penguin, pp. 235-237.

[110] Watson, P., (1981), *Twins: An Investigation into the Strange Coincidences in the Lives of Separated Twins,* London: Hutchinson, p. 56.

[111] Watson, 1981, p. 59.

[112] Watson, 1981, p. 58.

[113] Segal, 2012, p. 132.

[114] Bouchard, T. J., Jr., & McGue, M., (1990), Genetic and Rearing Environmental Influences on Adult Personality: An Analysis of Adopted Twins Reared Apart, *Journal of Personality, 58,* 263-292, pp. 266-267.

[115] Segal, 2012, pp. 25, 54, 275.

[116] Segal, 2012, p. 96.

[117] Segal, 2012, p. 45.

[118] Jackson, D., (1980, October), Reunion of Identical Twins, Raised Apart, Reveals Some Astonishing Similarities, *Smithsonian,* pp. 48-56.

[119] Shields, 1962, p. 94.

[120] Bouchard, Lykken, et al., 1990, p. 224.

[121] Bouchard, T. J., Jr., (1983), Do Environmental Similarities Explain the Similarity in Intelligence of Identical Twins Reared Apart? *Intelligence, 7,* 175-184, p. 183.

[122] Mukherjee, 2016, p. 381.

[123] Sapolsky, 2017, p. 235.

[124] Segal, 2012, p. 334.

[125] Plomin, R., (2004), *Nature and Nurture: An Introduction to Human Behavioral Genetics*, Belmont, CA: Thompson Wadsworth, pp. 42, 59.

[126] Kamin, 1974, p. 56. This quotation was part of an italicized sentence.

[127] McGue, M., & Bouchard, T. J., Jr., (1984), Adjustment of Twin Data for the Effects of Age and Sex, *Behavior Genetics, 14*, 325-343, p. 325.

[128] McGue et al., (1993), Personality Stability and Change in Early Adulthood: A Behavioral Genetic Analysis, *Developmental Psychology, 29*, 96-109, p. 107.

[129] Harris et al., (2016), Personality Stability From Age 14 to Age 77 Years, *Psychology and Aging, 31*, 862-874.

[130] Richard J. Rose, quoted in Cole, K. C., (1995), "Innumeracy," in R. Jacoby & N. Glauberman (Eds.), *The Bell Curve Debate: History, Documents, Opinions* (pp. 73-80), New York: Times Books, p. 75.

[131] Farber, 1981, p. 77.

[132] Rose, R. J., (1982), Separated Twins: Data and their Limits, [Review of the Book *Identical Twins Reared Apart: A Reanalysis*, by S. Farber], *Science, 215*, 959-960.

[133] Segal, 2012, p. 61.

[134] Cassill, K., (1982), *Twins: Nature's Amazing Mystery*, New York: Atheneum, p. 183.

[135] Dusek, V., (1987), Bewitching Science, *Science for the People, 19, (6)*, 19-22.

[136] McGue & Bouchard, 1989, pp. 22-23.

[137] McGue & Bouchard, 1989, p. 23.

[138] Eaves et al., (1989), *Genes, Culture, and Personality: An Empirical Approach*, London: Academic Press, p. 48.

[139] Bouchard, 1987, p. 65; Bouchard, 1998, p. 261.

[140] Bouchard, 1997, p. 145.

[141] Hirsch, 1997, p. 220.

[142] Shao et al., (2008), Genetic Architecture of Complex Traits: Large Phenotypic Effects and Pervasive Epistasis, *Proceedings of the National Academy of Sciences, 105*, 19910-19914, p. 19910.

[143] Hemani et al., (2014), Detection and Replication of Epistasis Influencing Transcription in Humans, *Nature, 508*, 249-253; Zuk et al., (2012), The Mystery of Missing Heritability: Genetic Interactions Create Phantom Heritability, *PNAS, 109*, 1193-1198.

[144] Wahlsten, D., & Gottlieb, G., (1997), "The Invalid Separation of Effects of Nature and Nurture: Lessons from Animal Experimentation," in R. Sternberg & E. Grigorenko (Eds.), *Intelligence, Heredity, and Environment* (pp. 163-192), New York: Cambridge University Press, p. 163.

[145] Johnson et al., (2007), Genetic and Environmental Influences on the Verbal-Perceptual-Image Rotation (VPR) Model of the Structure of Mental Abilities in the Minnesota Study of Twins Reared Apart, *Intelligence, 35,* 542-562, pp. 548–549.

[146] Taylor, 1980, p. 7.

[147] Bouchard, Lykken, et al., 1990, p. 227.

[148] Bouchard, Lykken, et al., 1990, p. 228.

[149] Bouchard, Lykken, et al., 1990, pp. 227-228.

[150] Bouchard, T. J., Jr., (2016), "Genes and Behavior: Nature via Nurture," in R. Sternberg et al. (Eds.), *Scientists Making a Difference: One Hundred Eminent Behavior and Brain Scientists Talk about Their Most Important Contributions* (pp. 73-76), New York: Cambridge University Press, pp. 73-74.

[151] Bouchard, Lykken, et al., 1990, p. 227. For a recent critical analysis of model fitting and its underlying assumption, see Richardson, K., (2017), *Genes, Brains, and Human Potential: The Science and Ideology of Intelligence*, New York: Columbia University Press, Chapter 2.

[152] Joseph, 2015, Chapter 7 and Appendix C.

[153] Segal, 2012, p. 102.

[154] Bouchard, Lykken, et al., 1990, p. 227.

[155] Bouchard, T. J., Jr., (2008), "Genes and Human Psychological Traits," in P. Carruthers et al, (Eds.), *The Innate Mind: Foundations and the Future* (Vol. 3), p. 12, doi: 10.1093/acprof:oso/9780195332834.001.0001

[156] Segal, 2012, p. 61.

[157] McCourt et al., (1999), Authoritarianism Revisited: Genetic and Environmental Influences in Twins Reared Apart and Together, *Personality and Individual Differences, 27,* 985-1014, p. 1001; Newman, D. L., Tellegen, A., & Bouchard, T. J., Jr., (1998), Individual Differences in Adult Ego Development: Sources of Influences in Twins Reared Apart, *Journal of Personality and Social Psychology, 74,* 985-995, p. 992. See also Kamin, L. J., & Goldberger, A. S., (2002), Twin Studies in Behavioral Research: A Skeptical View, *Theoretical Population Biology, 61,* 83-95.

[158] Segal, 2012, p. 172. The controversial concept of emergenesis and "emergenic traits" was first put forward by Lykken in 1982. According to Segal (2012, p. 172), "Emergenesis refers to genetically influenced traits that do not run in families. Emergenic traits are thought to emerge out of complex configurations of polymorphic genes that come together by chance in an individual."

[159] For examples of MISTRA researchers dismissing or attempting to explain away DZA group correlations that did not fit genetic predictions, see Hur, Y. M., Bouchard, T. J., Jr., & Eckert, E., (1998), Genetic and Environmental Influences on Self-Reported Diet: A Reared-Apart Twin Study, *Physiology & Behavior, 64,* 629-636, p. 636; Johnson et al., 2007, p. 551; McGue & Bouchard, 1989, pp. 23, 31; Segal, 2012, pp. 62, 222.

[160] Joseph, 2015, pp. 141-142.

[161] Bouchard, T. J., Jr., Segal, N. L., & Lykken, D. T., (1990), Genetic and Environmental Influences on Special Mental Abilities in a Sample of Twins Reared Apart, *Acta Geneticae Medicae et Gemellologiae, 39,* 193-206, p. 205.

[162] Bouchard, Lykken, et al., 1990, p. 223.

[163] Kolata, G., (1990, October 12th, National Ed.), Study Raises the Estimate of Inherited Intelligence, *New York Times*, p. A00022.

[164] Segal, 2012, p. 105.

[165] Segal, N. L., & Johnson, W., (2009), "Twin Studies of General Mental Ability," in Y. Kim (Ed.), *Handbook of Behavior Genetics* (pp. 81-99), New York: Springer, p. 85.

[166] Bouchard & McGue, 1990, p. 263.

[167] Bouchard, 1998, p. 262.

[168] Bouchard, 1993a, p. 57. In the early 1980s, Bouchard submitted a MISTRA IQ report for publication in *Science* based on 29 MZA twin pairs, in addition to an apparently non-modest sample of only *12* DZA pairs (Segal, 2012, p. 104).

[169] Bouchard, Lykken, et al., 1990, p. 224.

[170] Kamin & Goldberger, 2002, p. 87.

[171] Bouchard, T. J., Jr., (1994b), "Twin Studies of Intelligence," in R. Sternberg (Ed.), *Encyclopedia of Human Intelligence* (Vol. 2, pp. 1091-1095), New York: Macmillan, p. 1094.

[172] Bouchard, 1993b, p. 20.

[173] Pedersen et al., 1992, p. 350.

[174] Segal, 2012, p. 62.

[175] Segal, 2012, p. 336.

[176] Purcell, S., (2013), "Statistical Methods in Behavioral Genetics," in Plomin et al., (2013), *Behavioral Genetics* (6th ed.; pp. 357-411), New York: Worth Publishers, p. 383.

[177] Plomin et al., (1990), "Behavioral Genetics and Personality," in L. Pervin (Ed.), *Handbook of Personality: Theory and Research* (pp. 225-243), New York: Guilford, p. 235.

[178] See also Joseph, J., (2013b), "The Lost Study: A 1998 Adoption Study of Personality that Found No Genetic Relationship Between Birthparents and Their 240 Adopted-Away Biological Offspring," in R. Lerner & J. Benson (Eds.), *Advances in Child Development and Behavior, 45,* 93-124, San Diego: Elsevier.

[179] In the 2007 Johnson et al. study, the researchers did not state the number of opposite-sex pairs in their sample of 52 DZA pairs. Because the final MISTRA sample of 56 DZAs contained 18 opposite-sex pairs, the minimum number of opposite-sex pairs in the 2007 sample would have been 14, and the maximum number would have been 18.

[180] The Wechsler IQ correlations are taken from Segal, 2012, p. 286, based on unpublished figures given to her by Bouchard in 2009. The number of MZA (74) and DZA (52) pairs was reported in Segal, 2012, p. 284.

[181] The MISTRA Raven's Progressive Matrices IQ correlations are taken from Johnson et al., 2007, p. 552, Table 3, Test #21.

[182] Segal, 2017, pp. 31-32.

[183] Segal, 2012, pp. 286-287.

[184] Plomin, R., & DeFries, J. C., (1980), Genetics and Intelligence: Recent Data, *Intelligence, 4,* 15-24, p. 23.

[185] Bouchard, Lykken, et al., 1990, p. 227, Figure 1.

[186] Bouchard, Lykken, et al., 1990, p. 225.

[187] There is a relationship between a correlation coefficient and an average absolute difference. Assuming that the behavioral characteristic is normally distributed, the latter can be calculated using a formula based on the correlation coefficient and the standard deviation. See Plomin & DeFries, 1980, p. 22.

[188] Ridley, M., (2004), *The Agile Gene: How Nature Turns on Nurture* [Originally published as *Nature via Nurture*], New York: Perennial, p. 79.

[189] Johnson et al., 2007, p. 551.

[190] Johnson et al., 2007, p. 552.

[191] Richardson, K. (1998), *The Origins of Human Potential*, London: Routledge, p. 121.

[192] Juel-Nielsen, 1980, Part I, p. 75.

[193] Bouchard & McGue, 1990, pp. 280-281, Table 7.

[194] Segal, 2012, p. 99.

[195] Plomin et al., 2013, p. 273.

[196] Paul, A. M., (2004), *The Cult of Personality Testing*, New York: Free Press, xi.

[197] Blinkhorn, S., & Johnson, C., (1990), The Insignificance of Personality Testing, *Nature, 348,* 671-672, p. 671.

[198] Segal, 2012, p. 301.

[199] Morgeson et al., (2007), Reconsidering the Use of Personality Tests in Personnel Selection Contexts, *Personnel Psychology, 60,* 683-729, p. 720.

[200] Graham, J. R., (1987), *The MMPI: A Practical Guide* (2nd ed.), New York: Oxford University Press, p. 116.

[201] Kline, P., (1993), *The Handbook of Psychological Testing*, London: Routledge, p. 129.

[202] Kline, P., (1995), "A Critical Review of the Measurement of Personality and Intelligence," in D. Saklofske & M. Zeidner (Eds.), *International Handbook of Personality and Intelligence* (pp. 505-524), New York: Cambridge University Press, p. 514.

[203] Bouchard, T. J., Jr., (1994a), Genes, Environment, and Personality, *Science, 264,* 1700-1701, p. 1701.

[204] Plomin et al., (1998), Adoption Results for Self-Reported Personality: Evidence for Nonadditive Genetic Effects?, *Journal of Personality and Social Psychology, 75,* 211-218, p. 211.

[205] See Joseph, 2015, Appendix B; Joseph, 2013b.

[206] Joseph, 2013b.

[207] Joseph, J., (2016, March 6), "Bewitching Science" Revisited: Tales of Reunited Twins and the Genetics of Behavior, [Web log post, *Mad in America* "The Gene Illusion"].

[208] Rosen, C. M., (1987, September), The Eerie World of Reunited Twins, *Discover*, pp. 36-46, p. 39.

[209] Horgan, J., (1999), *The Undiscovered Mind*, New York: The Free Press, p. 145.

[210] Dusek, 1987.

[211] Lykken et al., (1992), Emergenesis: Genetic Traits that May Not Run in Families, *American Psychologist, 47,* 1565-1577, p. 1556.

[212] Wyatt et al., (1984), Natural Levels of Similarities between Identical Twins and Between Unrelated People, *The Skeptical Inquirer, 9,* 62-66, p. 64.

[213] Juel-Nielsen, 1980, Part I, p. 57.

[214] Chambers, 2017.

[215] Farber, 1981, pp. 273-274.

[216] Chambers, 2017, p. 174.

[217] Bouchard, T. J., Jr., (2014), Genes, Evolution and Intelligence, *Behavior Genetics, 44,* 549-577, p. 568.

[218] Plomin, R., & von Stumm, S., (2018), The New Genetics of Intelligence, *Nature Reviews Genetics, 19,* 148-159, p. 148.

[219] For more on Robert Plomin's long history of making behavioral gene discovery claims that were not replicated, see Joseph, 2015, Chapter 10.

[220] Richardson, 2017, p. 19.

[221] McGue, M., & Bouchard, T. J., Jr., (1998), Genetic and Environmental Influences on Human Behavioral Differences, *Annual Review of Neuroscience, 21,* 1-24, p. 18.

[222] Bouchard, T. J., Jr., & McGue, M., (2003), Genetic and Environmental Influences on Human Psychological Differences, *Journal of Neurobiology, 54,* 4-45, p. 5.

[223] Segal, 2012, p. 244, italics in original.

[224] Anonymous, 2014.

[225] Joyner, M. J., Paneth, N., & Ioannidis, J. P. A., (2016), What Happens When Underperforming Big Ideas in Research Become Entrenched?, *JAMA, 316,* 1355-1356, p. 1356.