

A Reevaluation of the 1990 “Minnesota Study of Twins Reared Apart” IQ Study

Jay Joseph

Private practice, Oakland, CA, USA

Keywords

Twin studies · IQ · Intelligence · Behavioral genetics · Reared-apart twins · Heritability · Minnesota twin study

Abstract

In 1990, Thomas J. Bouchard Jr. and colleagues published the widely cited 1990 “Minnesota Study of Twins Reared Apart” (MISTRA) *Science* IQ study. To arrive at the conclusion that “IQ is strongly affected by genetic factors,” Bouchard and colleagues omitted their control group reared-apart dizygotic twin (DZA) IQ score correlations. Near-full-sample correlations published after the study’s 2000 end point show that the reared-apart monozygotic twin (MZA) and DZA group IQ correlations did not differ at a statistically significant level, suggesting that the study failed the first step in determining that IQ scores are influenced by heredity. After bypassing the model-fitting technique they used in most non-IQ MISTRA studies, the researchers assumed that the MZA group IQ score correlation alone “directly estimates heritability.” This method was based on unsupported assumptions by the researchers, and they largely overlooked the confounding influence of cohort effects. Bouchard and colleagues then decided to count most environmental influences they did recognize as genetic influences. I conclude that the MISTRA IQ study failed to discover genetic influenc-

es on IQ scores and cognitive ability across the studied population, and that the study should be evaluated in the context of psychology’s replication problem.

© 2022 S. Karger AG, Basel

Problems in IQ Studies Based on Reared-Apart Twins

In this article, I will examine the claim that the widely cited “Minnesota Study of Twins Reared Apart” (MISTRA) produced evidence that genetic factors play a role in causing differences in human intelligence (cognitive ability), as supposedly measured by IQ tests and heritability estimates. The MISTRA is one of five “twins reared apart” (TRA) studies that published IQ correlations for what the investigators regarded as reared-apart twins (Bouchard et al., 1990a; Juel-Nielsen, 1965/1980; Newman et al., 1937; Pedersen et al., 1992; Shields, 1962). In a widely covered controversy, British psychologist Cyril Burt’s IQ TRA study was discredited in the 1970s on suspicions of fraud, and Burt’s publications are no longer part of the TRA IQ literature (Bouchard & McGue, 1981; Hearnshaw, 1979; Kamin, 1974; Lerner, 2018; Tucker, 1997).

There are two types of twin pairs: MZ (monozygotic, identical), and DZ (dizygotic, fraternal). MZ pairs are

said to share 100% of their segregating genes, whereas (like ordinary siblings) DZ pairs are said to share only 50% on average. Reared-apart MZ twins are known as MZA pairs. Reared-apart DZ twins are known as DZA pairs. Recent evidence, however, suggests that the long-running assumption that MZ pairs are genetically identical to each other might not be true (Charney, 2012; Jonsson et al., 2021).

Contemporary TRA researchers assume that if the members of an MZA pair grow up apart in “uncorrelated” family environments, all behavioral resemblance between them is caused by their purported 100% genetic similarity (Bouchard, 1984; Pedersen et al., 1992; Segal, 2012). In 2004, leading behavioral geneticist and twin researcher Robert Plomin (2004) 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” (p. 59). “How else,” Plomin asked, “can we explain” the behavioral correlations of “identical twins adopted apart into uncorrelated environments” (p. 42). I argue in this paper that this assumption is not supported by the evidence.

Criticism of TRA Studies

Over the years, commentators have described many biases and problems found in TRA studies, especially those that, like the MISTRA, Newman et al. (1937), and Shields (1962) studies, were based on volunteer twins (Joseph, 2015, chapter 3). Psychologist Leon Kamin was a pioneering analyst of TRA research (Joseph, 2018; Kamin, 1974), and psychologist Susan Farber (1981) published an exhaustive critical analysis of the three TRA studies conducted prior to the MISTRA.

Detailed case history information was provided only by the authors of the first three TRA studies (Juel-Nielsen, 1965/1980; Newman et al., 1937; Shields, 1962). I have compiled some characteristics of these studies’ combined sample of 75 MZA pairs: (a) in 25/75 (33%) of the pairs, twins were separated at 12 months of age or later; (b) in 56/75 (75%) of the pairs, twins had contact with each other while growing up; (c) in 42/75 (56%) of the pairs, one or both twins were placed with a family member; and (d) in 17/75 (23%) of the pairs, twins lived together for a period of at least 12 months after separation, or grew up next door to each other. (For specific examples of twins’ late/inadequate separation and prestudy contact found in these 75 case histories, see Joseph, 2015, Tables 2.1–2.3).

For these and other reasons, Farber accurately concluded, “the assumption that this is a sample of reared-apart twins is incorrect. Most were reared apart only partially, and almost all had contact with each other by the time they were studied” (Farber, 1981, p. 19).

Additional problems, biases, and potential environmental confounds described by critics of TRA research include (a) the fact that twins share a prenatal environment (Farber, 1981); (b) the selective placement of twins and the lack of random assignment (Fancher, 1985; Farber, 1981; Kamin, 1974; Taylor, 1980); (c) that twins usually grew up in similar socioeconomic and cultural environments (Farber, 1981; Kamin, 1974; Rutter, 2006; Taylor, 1980); (d) methodological bias in favor of the recruitment of the most behaviorally similar segment of the total population of MZA pairs (Fancher, 1985; Farber, 1981; Kamin, 1974; Lerner, 2018; Rutter, 2006); (e) that cohort influences, in addition to twins’ common age and common sex, increase twins’ IQ correlations and behavioral resemblance in general (Kamin, 1974; Rose, 1982; Wyatt, 1993); (f) the potential unreliability of twins’ accounts of their degree of separation/contact and behavioral similarity (Joseph, 2015; Kamin, 1974; Kamin & Goldberger, 2002); (g) researcher confirmation bias (conscious or unconscious) in favor of genetic interpretations of the results (Joseph, 2015; Kamin, 1974); (h) problems with the validity of the tests administered to the twins (Kamin, 1974; Rutter, 2006; Taylor, 1980); (i) that the results might not apply (generalize) to the nontwin population (Cropanzano & James, 1990; Farber, 1981; Lerner, 2018); and (j) that MZAs will elicit more similar treatment from their social environments due to their substantial physical resemblance (Billings et al., 1992; Cropanzano & James, 1990; Ford, 1993).

In 1985, psychologist and historian Raymond Fancher described what he believed a “definitive” and “ideal” 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 (Fancher, 1985, p. 165, emphasis in original).

The MISTRA did not come close to meeting this standard. As the study’s initiator and leader Thomas J.

Bouchard Jr. recognized in relation to the published TRA studies, including his own,

Twins are not literally separated at birth, randomly assigned to homes, and evaluated as adults prior to any social contact. Such procedures would be necessary if a real experiment were being conducted, but such an experiment would be unethical. (Bouchard, 1993, p. 56)

Because a properly performed TRA study would be unethical, or is virtually impossible to accomplish, does not mean that science, and the general public informed by media accounts, must accept the results of a genetic study if it is based on a research design unable to control for potential environmental confounds.

I will now focus on the MISTRA IQ study, and on how the researchers arrived at their conclusions in favor of the “strong heritability” of IQ (Bouchard et al. 1990a, p. 223) – a conclusion I will challenge. The story I will tell about this famous study differs greatly from the story usually told by journalists, textbook authors, leading behavioral geneticists, and other authoritative authors and sources (references below).

The Minnesota Study of Twins Reared Apart IQ Study

“The immediate causes of most psychological variations are probably environmental in nature,” Nancy Segal wrote in *Born Together – Reared Apart* (2012, p. 111).

Psychologist and MISTRA researcher Nancy L. Segal summarized the MISTRA findings by concluding that most human psychological variations are “probably environmental in nature.” But because she and her colleagues concluded in favor of important genetic influences on variations in IQ, personality, and other areas of human behavior, she must have followed this statement with a “but,” or with a “however.” Indeed, she did.

The MISTRA IQ results were published in a 1990 edition of *Science*, one of the world’s leading scientific journals (Bouchard et al., 1990a; according to Google Scholar, as of January 2022, the article had been cited over 2,200 times). Although the MISTRA researchers published findings related to personality and other areas of human behavior, findings and conclusions about their index of cognitive ability (IQ) are the focus of the present review. As the researchers recognized, “IQ has been at the center of the nature-nurture debate” (Bouchard et al., 1990a, p. 224).

Bouchard’s MISTRA colleagues included Segal, psychologists David T. Lykken, Auke Tellegen, Matt McGue, and others. The study was conducted between 1979 and 2000. Additional publications based on the MISTRA data

appeared after the study’s 2000 end date (e.g., Johnson et al., 2007). In 2012, Segal published *Born Together – Reared Apart: The Landmark Minnesota Twin Study*. This book provided a detailed description of the study and won the American Psychological Association’s (Division 1) 2013 “William James Book Award.” The following year, the American Psychological Foundation presented Bouchard with its 2014 “Gold Medal Award for Lifetime Achievement in the Science of Psychology” for his MISTRA work (Gold medal, 2014).

The 1990 *Science* 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 (Segal, 2012). In Bouchard’s view, this article “legitimated the study” (quoted in Segal, 2012, p. 104).

The MISTRA Sample

The MISTRA was not a longitudinal twin study, and like most behavioral genetic investigations it was based on participants who grew up in environments the researchers were unable to design, control, or observe. For this reason, they had to make assumptions about twins’ unobserved environments, and their conclusions depended on the validity of these assumptions.

The MISTRA reared-apart twin sample at the study’s 2000 end point consisted of 81 MZA and 56 DZA pairs, 45% of whom were recruited from the USA. The rest grew up outside of the USA, and 32% came from England (Segal, 2012). Like the studies by Newman et al. (1937) and Shields (1962) that preceded it, the MISTRA sample was a “collection of cases” based mainly on volunteer twin participants who responded to media appeals or media attention (Segal, 2012, p. 36).

At the time of the 1990 *Science* publication, the MISTRA sample consisted of 56 MZA and 30 DZA pairs. The researchers provided very little information about the twins they studied. Some characteristics of the 1990 MZA sample they chose to publish included “mean age” = 41 years, mean “time spent together prior to separation” = 5.1 months, mean “time apart to first reunion” = 30 years, mean “total contact time” = 112.5 weeks (with a range of 1–1,233 weeks), and a mean IQ of 108.1 (Bouchard et al., 1990a, p. 224). The researchers used the intraclass correlation method to calculate twin IQ and other psychological test score correlations (discussed in Segal, 2012).

Table 1. Reported and nonreported correlations in the 1990 MISTRA *Science* Article

Measure	MZA	MZT	DZA control group
<i>WAIS full-scale IQ</i>	0.69	0.88	not reported
<i>Raven, Mill-Hill composite IQ</i>	0.78	0.76	not reported
<i>First principal component IQ</i>	0.78	"not available"	not reported
Hawaii Special Mental Abilities	0.45	"not available"	not reported
Comprehensive Special Mental Abilities	0.48	"not available"	not reported
Multidimensional Personality Questionnaire	0.50	0.49	not reported
California Psychological Inventory (personality)	0.48	0.49	not reported
Strong Campbell Interest Inventory	0.39	0.48	not reported
Religiosity Scales	0.49	0.51	not reported
MPQ Traditionalism Scale	0.53	0.50	not reported

IQ measures are italicized. Intraclass correlations from Bouchard et al. (1990a, p. 226), Table 4. DZA, DZ twins reared apart; MZA, MZ twins reared apart; MZT, MZ twins reared together; WAIS, Wechsler Adult Intelligence Scale; MPQ, Multidimensional Personality Questionnaire. The 1990 MISTRA *Science* sample consisted of 56 MZA and 30 DZA pairs (p. 223). The MISTRA-reported MZA IQ correlations were based on fewer pairs, ranging from 42 to 48 pairs. "Not available" status was reported by the researchers. No DZA correlations or results of any kind were reported in the 1990 MISTRA *Science* article.

Impact of the Study

"Studies of general intelligence," Segal wrote, "were a mainstay of the MISTRA's research program" (2012, p. 284). Early on, Bouchard saw "intelligence as the major mental ability" (Bouchard, 1981, p. 22, emphasis in original) and later noted that the "study of IQ is paradigmatic of human behavior genetic research" (Bouchard et al., 1990a, p. 224). The MISTRA IQ study has had a major impact on the nature-nurture debate in support of nature, and its authors' conclusions in support of a major role for genetic influences on behavioral differences have been cited, and sometimes celebrated, in textbooks (e.g., Bernstein et al., 2000; Hewstone et al., 2005; Myers, 2005; Plomin et al., 2013), in popular works (e.g., deBoer, 2020; Harris, 1998; Mukherjee, 2016; Pinker, 2002; Ridley, 2004; Sapolsky, 2017; Wright, 1997), and in the mass media (e.g., Angle & Neimark, 1997; Begley, 1987; Brody, 2018; Lewis, 2019).

In *Behave: The Biology of Humans at Our Best and Worst*, neuroscientist Robert Sapolsky (2017) referred enthusiastically to the MISTRA as "the most wonderful, amazing, like, totally awesome thing ever in behavior genetics, started by Thomas Bouchard of the University of Minnesota" (p. 235). A critical commentator described the study as "arguably, the most famous social science project in the last quarter of the twentieth century" (Tucker, 2013, p. 337). Segal (2012) believed that Bouchard and the study "forever changed the way people think about

the roots of human behavior" (p. 400). The "results from the *Science* paper," Segal (2012) wrote, "appeared in hundreds of newspapers, magazines, and broadcasts across the country and around the world" (p. 112).

Reported MISTRA IQ Findings

As Segal (2012) wrote, the MISTRA "*Science* IQ paper," which presented the "long-awaited IQ data" (p. 104), was published in 1990. The three IQ measures used in the study were the "Wechsler Adult Intelligence Scale" (WAIS; Wechsler, 1955), the "Raven's Progressive Matrices/Mill-Hill Vocabulary Scale composite" (Raven, 1986), and the "First Principal Component of Special Mental Abilities," or FPC (Bouchard et al., 1990a, p. 226). The researchers later published MZA and DZA subsample cognitive ability correlations drawn from larger samples than were available in 1990. These subsample correlations were reported in the context of MISTRA studies of "adult ego development" (Newman et al., 1998), and "right-wing authoritarianism" (McCourt et al., 1999). Selected MISTRA 1990 *Science* results are seen in Table 1.

Based on the 1990 *Science* findings seen in Table 1, Bouchard et al. estimated the IQ "variance due to genetic differences" (heritability) at "about 70%," and concluded that "genetic factors exert a pronounced and pervasive influence on behavioral variability" (pp. 223–224). I discuss below how they arrived at these conclusions.

Bouchard's conclusions matched his previously held IQ-hereditarian view that genes play a major role in determining IQ scores. IQ hereditarianism is the belief that heredity is the dominant factor influencing individual and sometimes group differences in intelligence. By his own account, Bouchard had been an early supporter and teacher of psychologist Arthur Jensen's (1969) IQ-hereditarian theories, and because of this role he drew protests from University of Minnesota students in the early 1970s (Holden & Bouchard, 2009).

In a pre-MISTRA 1976 chapter entitled "Genetic Factors in Intelligence," Bouchard argued in IQ-hereditarian fashion that "human intelligence," as supposedly measured 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" (Bouchard, 1976, p. 193). Two decades later, Bouchard (1995, p. 417) and Lykken (1995, pp. 216–217) endorsed Richard Herrnstein and Charles Murray's conclusion in *The Bell Curve* that "both genes and the environment have something to do with racial differences" in IQ scores (Herrnstein & Murray, 1994, p. 311).

How the MISTRA Researchers Concluded in Favor of 70% IQ Heritability

The Science IQ Study Abstract

Before looking at problem areas in the main body of the 1990 MISTRA *Science* publication, and with an understanding that an abstract is necessarily noncomprehensive, it is useful to look at the article's full abstract as a frame for my comments about it, which will follow.

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. (Bouchard et al. 1990a, p. 223)

In this abstract, Bouchard et al., who framed their conclusions in the context of previous behavioral genetic research, did not say how they determined that "about 70% of the variance in IQ was found to be associated with genetic variation," or how they concluded in favor of the "strong heritability of most psychological traits." In addition, they did not clearly state what conclusions followed from the finding that their MZA and MZT (reared-together MZ twins) behavioral correlations were similar. Because this result was the only twin-based finding reported in the abstract, a reader might conclude that this finding led to the researchers' "strong heritability" conclusion, and to their 70% IQ heritability estimate. As we will see, the researchers did not compare their MZA and MZT correlations for the purpose of establishing and calculating heritability. Bouchard et al. believed that it is "a plausible hypothesis" that "genetic differences affect psychological differences largely indirectly," which suggests that *non-genetic* differences affect psychological differences largely directly, seemingly contradicting their "strong heritability" conclusion. The abstract stated that the study tested "more than 100 sets of reared-apart twins or triplets," but the article that followed reported IQ correlations for less than half of these sets (WAIS MZA $N = 48$ pairs, Raven/Mill-Hill MZA $N = 42$ pairs, FPC MZA $N = 43$ pairs; Bouchard et al., 1990a, p. 226; no DZA correlations were reported). Finally, although the abstract stated that the study's findings do not "detract from the value or importance of parenting," in the main body of the article, the researchers concluded that their findings suggested that "common rearing enhances familial resemblance during adulthood only slightly and on relatively few behavioral dimensions" (p. 227).

DZAs Were the MISTRA Designated Control Group

Bouchard decided at the beginning of the study in 1979 to use DZA pairs to form the MISTRA control group, which he planned to compare to the MZA experimental group to determine whether genetic factors influence human behavioral differences. According to Nancy Segal (2012), "The MISTRA was ... the first study of its kind to include both MZA and DZA twin pairs. Bouchard's decision to use DZA twins as controls was made in a very early [1979] memo to the 'Twin Research Team.' This was an important methodological improvement over past projects" (p. 12). In a 1986 publication, appearing 4 years prior to the publication of their 1990 *Science* IQ study, Bouchard, Segal, and colleagues affirmed their plan to use

DZAs as the MISTRA control group to test genetic versus “competing” nongenetic explanations of TRA study results:

Our study is the first to have included a control group of dizygotic twins reared apart (DZA). ... 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. (Bouchard et al., 1986, p. 300)

Bouchard and colleagues continued this theme in the opening paragraph of their 1990 *Science* article. “MZA and DZA twin pairs,” they wrote, “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” (Bouchard et al., 1990a, p. 223). However, although the 1990 *Science* article was the main MISTRA IQ publication, and although the 1990 sample consisted of 56 MZA and 30 DZA pairs, this “fascinating” and “most powerful” “experiment of nature” did not appear in the article. No DZA results of any kind were reported in the 1990 *Science* article, due to what Bouchard and colleagues said were “space limitations and the smaller size of the DZA sample (30 sets).” They believed that the “smaller” DZA sample justified their decision to “focus on the MZA data (56 sets)” (Bouchard et al., 1990a, p. 223). The researchers arrived at this decision in 1990 despite the fact that, a few years earlier, Bouchard had submitted a MISTRA IQ study for publication in *Science* that included data from a sample of 29 MZA and only 12 DZA pairs (Segal, 2012, p. 104).

Interpreting the IQ Study Results

Bouchard and colleagues did not compare the MZA and MZT correlations seen in Table 1 to assess the role of genetics. They used this comparison only to determine whether being reared together in the same family environment leads to behavioral similarity (Bouchard et al., 1990a, pp. 226–227). Their conclusion in favor of the “strong heritability” of IQ (70%) was based on the MZA IQ correlations alone (WAIS, Raven/Mill Hill, FPC), based on the claim that because MZ twins share a 100% genetic identity, and “under the assumption of no environmental similarity,” the MZA correlation “directly estimated” IQ heritability (Bouchard et al., 1990a, p. 224). Segal restated this commonly repeated behavioral genetic assumption in 2012: “The MZA twin correlation directly estimates heritability” (p. 190). Focusing on their MZA

data alone to determine heritability allowed the researchers to bypass/omit their DZA control group correlations.

Although the researchers omitted their DZA correlations “due to space limitations and the smaller size of the DZA sample (30 sets),” in *non-IQ* MISTRA studies appearing around the same time, they published their MZA and full-sample DZA correlations based on similar DZA sample sizes. Three examples are the MISTRA 1988 personality study, where the researchers published correlations based on a sample of 27 DZA pairs (Tellegen et al., 1988), a 1990 MISTRA study of “religious interests, attitudes, and values” based on 31 DZA pairs (Waller et al., 1990), and a 1990 MISTRA “special mental abilities” study based on 25 DZA pairs (Bouchard et al., 1990b).

The MISTRA Model-Fitting Technique Was Not Used in the 1990 *Science* Article

When assessing the potential role of genetic factors, according to Segal (2012), “most of the MISTRA’s heritability analyses used biometrical modeling techniques as well as intraclass correlations” (p. 62). Although prior to the IQ study Bouchard had written that “a model-fitting approach to family data ... is far superior to the simple calculation of heritabilities” (Bouchard, 1984, p. 150), and after the IQ study that “formal testing of substantive hypotheses via model fitting has now become the norm in behavior-genetic research” (Bouchard, 1997, p. 129), the 1990 MISTRA IQ study did not use the model-fitting technique.

Biometrical model fitting, a technique developed in the 1970s (Jinks & Fulker, 1970), is widely used in behavioral genetic research. The technique attempts to partition sources of behavioral variation using the “ACE” model, which supposedly partitions (A) additive genetic effects, (C) shared family environment, and (E) unique (unshared) environment (Purcell, 2013). Heritability estimates are based on “A” and range from 0.0 to 1.0, or 0 to 100%. Nevertheless, the validity of heritability estimates, and the claim that sources of behavioral variation are independent of each other (do not interact) and can be partitioned have been challenged for decades (Keller, 2010; Lerner, 2018; Lewontin, 1974; Moore & Shenk, 2016). The validity of psychometric IQ testing and the “general intelligence” concept (“g”) have also been called into question (Fischer et al., 1996; Gould, 1981; Lewontin et al., 1984; Mensh & Mensh, 1991; Richardson, 2017).

Because the MISTRA researchers decided against using their model-fitting procedure in their IQ study, I refer

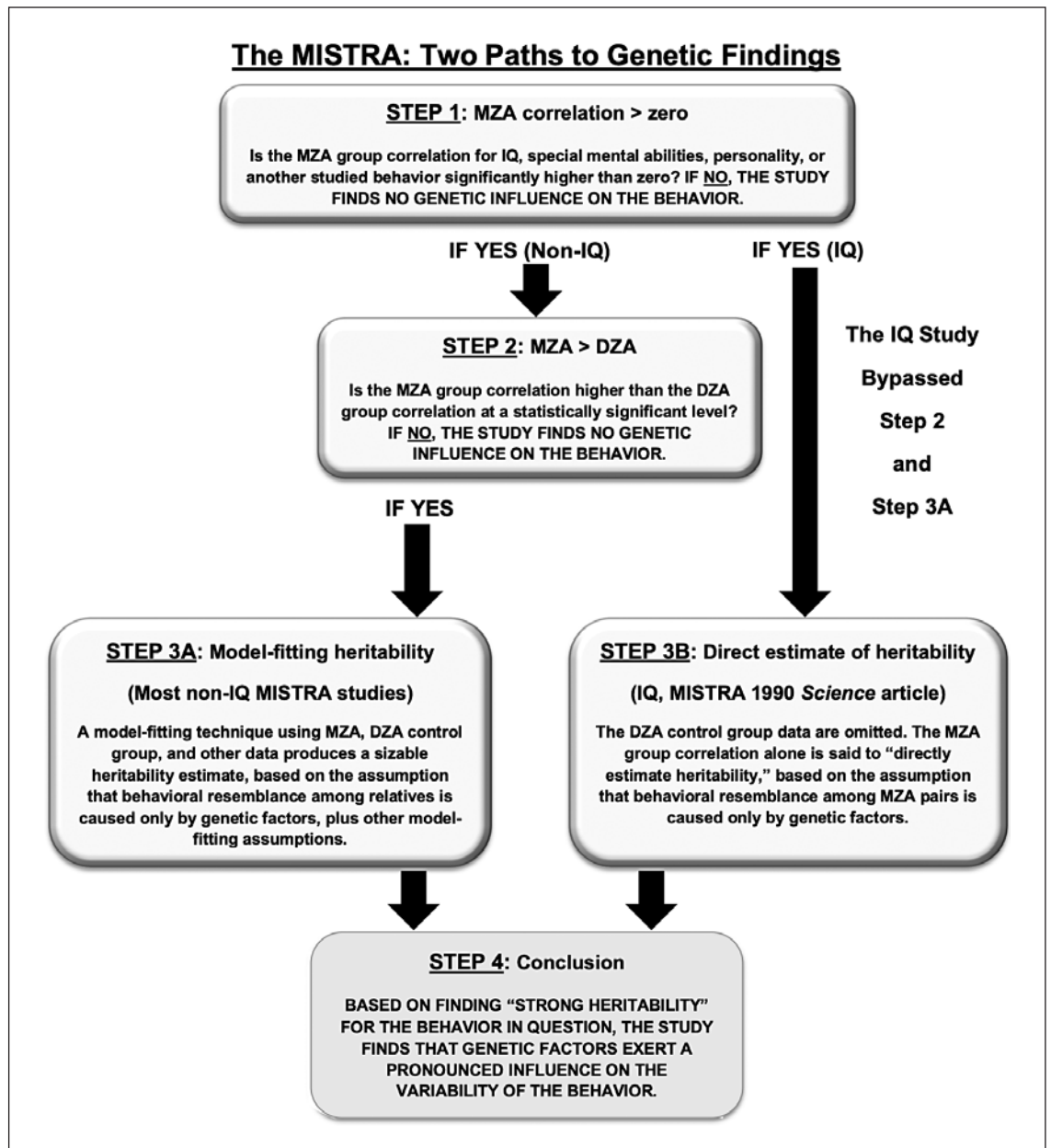


Fig. 1. Diagram of the two paths used to arrive at conclusions in favor of genetics.

readers to publications by authors who questioned basic model-fitting assumptions (Goldberger, 1979; Joseph, 2015; Lerner, 2018; Lewontin, 1974; Taylor, 1980). Still, it is worth noting that McGue and Bouchard recognized, in relation to the MISTRA model-fitting assumptions, that “several of these assumptions are likely not to hold for cognitive abilities” (McGue & Bouchard, 1989, p. 23).

Two Paths to Genetic Conclusions

As we have seen, the two main methods the MISTRA researchers used to estimate the heritability of psychological characteristics (traits) were: (a) biometrical model-fitting techniques, which they used in most of their non-IQ publications (e.g., Bouchard & McGue, 1990; Bouchard et al., 1990b; DiLalla et al., 1996; Hur & Bouchard, 1995; Keller et al., 1992; Koenig et al., 2005; Newman

et al., 1998; Tellegen et al., 1988; Waller et al., 1990) or (b) assessing the MZA correlation alone to “directly estimate” heritability (Arvey et al., 1989; Bouchard et al., 1990a). Figure 1 is a diagram I created to show the two paths the researchers took to arrive at their conclusions in favor of genetics.

What we might call the MISTRA “IQ path,” used in the 1990 *Science* article, is seen on the right side of Figure 1. Creating this path allowed the researchers to bypass their control group DZA correlations (step 2) and their model-fitting procedures (step 3A), and to calculate IQ heritability based on the MZA correlation alone (step 3B). What we might call the MISTRA “non-IQ path,” used in most non-IQ MISTRA studies, is seen on the left side of Figure 1. When using this path, the researchers based their conclusions on the MZA and control group DZA correlations, and on model-fitting procedures.

Like most psychological research of that era, the investigators’ decisions to create two paths rather than one, and about which path they would follow in a given publication, were made behind the scenes with the data and correlations at hand (more on this point later).

The MZA Correlation Must Significantly Exceed the DZA Correlation

As stated in a leading behavioral genetics textbook discussion of model fitting, “simple comparisons between twin correlations can indicate whether genetic influences are important for a trait. This is the *important first question* [emphasis added] that any quantitative genetic analysis must ask” (Purcell, 2013, p. 383). 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* [emphasis added] (unless assortative mating approaches unity), and model-fitting analyses must come to that conclusion or there is something wrong with the model. (Plomin et al., 1990, p. 235)

As Plomin stressed, a model-fitting analysis finding genetic influences and estimating heritability would be wrong if the MZT correlation does not (significantly) exceed the DZT (reared-together DZ twins) correlation, a conclusion that applies to MZA-DZA compari-

sons as well (Fig. 1, step 2). If the MZ correlation is not higher than the DZ correlation, Plomin wrote, “there is no genetic influence” on the behavioral characteristic (trait).

In *Born Together – Reared Apart*, Segal (2012) confirmed that comparing the MZA and DZA group correlations is an “important first step” in determining “whether or not” genetic factors influence a trait: “The simple comparison of the MZ (or MZA) and DZ (or DZA) intraclass correlations is an *important first step* [emphasis added] in behavioral-genetic analysis because this demonstrates *whether or not* [emphasis added] there is genetic influence on the trait” (p. 62). Elsewhere in the book, Segal (2012) wrote: “Genetic effects are shown *if* [emphasis added] the correlation for MZ or MZA twins exceeds the correlation for DZ or DZA twins” (p. 336).

The MISTRA MZA and DZA IQ Correlations Did Not Differ at a Statistically Significant Level

To the best of my knowledge, Bouchard and colleagues never published their full-sample DZA control group IQ correlations (Joseph, 2015, chapter 6; Kamin & Goldberger, 2002). A plausible explanation for why they chose not to is that the results would have shown that the MZA group correlation was not significantly higher than the DZA group correlation for any of the three IQ measures used in the 1990 IQ study.

The MISTRA near-full-sample MZA and DZA WAIS IQ correlations were published by Segal in 2012, almost as an afterthought towards the end of *Born Together – Reared Apart*. Segal obtained these correlations “from Bouchard when I visited Minnesota in October 2009 ... I imagine that other readers wanted them as well” (pp. 285–286). In a 2007 MISTRA study of the “Verbal-Perceptual-Image Rotation (VPR) model of the structure of mental abilities,” the researchers published their near-full-sample Raven’s Progressive Matrices IQ MZA and DZA correlations in a table, without discussing the meaning of these correlations (Johnson et al., 2007, p. 552, Table 3, Test 21).

The reported samples from both Segal (2012) and Johnson et al. (2007) were based on 74 MZA and 52 DZA pairs, a few pairs short of the final MISTRA sample of 81 MZA and 56 DZA pairs. Because neither report stated whether these groups constituted the MISTRA full sample minus twin pairs who did not complete the IQ test, or whether some pairs and their IQ scores were not included in the samples, for the purpose of the present review I refer to the

Table 2. Near-full-sample MISTRA IQ correlations at the study's end: MZA versus DZA twin pairs

	MZA pairs (experimental group)	DZA pairs (control group)	Probability value
Wechsler (WAIS) IQ correlations	0.62 (74 pairs)	0.50 (52 pairs)	$p = 0.17$ Not statistically significant at the 0.05 level
Raven's Progressive Matrices IQ correlations	0.55 (74 pairs)	0.42 (52 pairs)	$p = 0.18$ Not statistically significant at the 0.05 level

Intraclass correlations. One-tailed probability. The final 2000 MISTRA full sample consisted of 81 MZA and 56 DZA pairs. Based on calculations made at the VassarStats website (<http://vassarstats.net/rdiff.html>). DZA, DZ twins reared apart; MZA, MZ twins reared apart; p , one-tailed probability; WAIS, Wechsler Adult Intelligence Scale. Sources: Wechsler (WAIS) correlations from Segal (2012, p. 286), based on the number of pairs reported on p. 284; Raven correlations from Johnson et al. (2007, p. 552), based on the number of pairs reported on p. 545. The DZA sample contained 18 opposite-sex pairs (Segal, 2012, p. 42). The DZA group correlation for the MISTRA "First Principal Component of Special Mental Abilities" measure, which was the third of three MISTRA IQ measures, has never been published.

Segal (2012) and Johnson et al. (2007) correlations as being based on the near-full-sample MISTRA data.

The results reported by Segal and Johnson et al. are seen in Table 2. To the best of my knowledge the researchers never published their DZA FPC correlations, and no Minnesota FPC DZA correlations were reported in Segal's (2012) book.

We see in Table 2 that the MISTRA IQ study could not get past what Segal called the "important first step" (Fig. 1, step 2) of discovering genetic influences, because there was no statistically significant difference between the MISTRA near-full-sample MZA and DZA correlations for either the WAIS or the Raven IQ tests. These results indicate that from a statistical and IQ research standpoint, the MISTRA findings showed that MZAs' greater genetic similarity versus DZAs (100 vs. 50%) did not result in significant differences between the MZA and DZA correlations. These results lead to a conclusion that the study found no evidence that differences in IQ scores across the studied population are consistent with differences in genetic factors. Instead, all IQ differences across the studied population are consistent with differences in nongenetic factors alone. In behavioral genetic terms, the study did not find any reliable evidence that IQ is heritable.

Nevertheless, Bouchard and colleagues concluded that "general intelligence or IQ is strongly affected by genetic factors" (Bouchard et al., 1990a, p. 227). Before examining additional reasons why this conclusion was wrong, in the following section I will argue that the MISTRA IQ study should be evaluated in the context of problem areas in the psychological research and publication process that have come to light in recent years.

Questionable Research Practices

In a 1998 article, Bouchard recognized that although the DZA correlation "is reported as a control variable" in the MISTRA, the study's full-sample 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" (Bouchard, 1998, p. 262; on page 263 of this article, Bouchard reported a MISTRA 28-pair DZA subsample WAIS correlation of 0.47). When TRA study critic Leon Kamin asked Bouchard in 1997 to share the MISTRA full-sample IQ correlations, Bouchard responded, "I can't pass on the IQ results for our MZA's or DZA's [*sic*] because I have not published them yet. Indeed, I have not even calculated them" (Kamin & Goldberger, 2002, p. 87; see also Joseph, 2015, chapter 6). Perhaps these closely guarded "not even calculated" high DZA group correlations would have enabled Kamin and others to show that the MISTRA results produced no evidence in support of genetic influences on any of the three MISTRA IQ measures.

p -Hacking

In his book *The Seven Deadly Sins of Psychology: A Manifesto for Reforming the Culture of Scientific Practice*, cognitive neuroscientist Chris Chambers (2017) described several major problem areas in the research/publication process in psychology and other fields. One of these is "hidden flexibility," which refers to researchers' behind-the-scenes ability to change various aspects of their study after reviewing their data, but before submitting their paper for publication. Under the long-running

psychology publication system, undetected “questionable research practices” appear to be common (John et al., 2012), and the field has been shaken in recent years by study retractions (for example, the work of the famous psychologist H.J. Eysenck; O’Grady, 2020), “*p*-hacking” (Head et al., 2015), fabricated research (Carey, 2011), and a replication crisis (Baker, 2015; Open Science Collaboration, 2015). These problems were also discussed at length by psychologist Stuart Ritchie in his 2020 book *Science Fictions: How Fraud, Bias, Negligence, and Hype Undermine the Search for Truth*.

Both Chambers and Ritchie called for the establishment of research “pre-registration,” where investigators would have the option or be required to submit their research rationale, hypotheses, design and analytic strategy, and planned data collection stop point to a journal for peer review before collecting and analyzing data (see also Joseph & Baldwin, 2000).

p-hacking is the practice of consciously or unconsciously manipulating data to produce results that fall below the 0.05 level of statistical significance. As described by Megan Head et al. in 2015, the “widespread” practice of *p*-hacking occurs “when researchers collect or select data or statistical analyses until nonsignificant results become significant” (Head et al., 2015, p. 1), which is sometimes achieved by continuing to collect data past the planned stop point if significant comparisons are not found. Another *p*-hacking method “occurs when researchers try out several statistical analyses and/or data eligibility specifications and then selectively report those that produce significant results” (Head et al., 2015, p. 1).

p-Hacking in the 1990 *Science* IQ Study

In the area of IQ, the MISTRA researchers appear to have engaged in *p*-hacking (a) when they failed to publish and assess their control group DZA IQ correlations at the 1990 *Science* study stop point; and (b) when in the same article they selectively reported a method that produced statistically significant results, while failing to report the results of the planned method (MZA-DZA comparison and/or model fitting) which, the evidence suggests, produced statistically nonsignificant results.

It is also possible that the MISTRA researchers intended to continue withholding their DZA correlations from publication or independent review until the sample became large enough to nudge the MZA-DZA comparison under the critical 0.05 level of statistical signif-

icance. Eight years after the publication of the MISTRA IQ study, Bouchard (1998) claimed that the MISTRA full-sample DZA IQ correlations “have not yet been fully analyzed.” The following year, Segal (1999) wrote that in this 1998 article, “Bouchard cautioned that the Minnesota [IQ] data are preliminary and require further analysis” (p. 136). Neither Segal nor Bouchard, however, provided a valid reason *why* the Minnesota data were preliminary and required further analysis. Questionable research practices of this type can occur when researchers are not required to adhere to a stated data collection stop point, which would be established and documented in a pre-registered study. In a different context, Ritchie (2020) noted that “not setting the sample size beforehand allows ... researchers to continue collecting data and testing it, collecting data and testing it, again and again in an open-ended way until they get their desired $p < 0.05$ ” (p. 134).

Psychology’s “Replication Crisis”

Although undoubtedly many researchers in psychology did not engage in such practices, we have seen that the MISTRA was performed and published in an era in which many of the research practices leading to psychology’s “replication crisis” were common. Publications appearing in peer-reviewed psychology journals were often presented as neatly packaged articles, with little prior record of researchers’ intended methods, assumptions, definitions, comparisons, and decision-making processes (Chambers, 2017; Joseph & Baldwin, 2000).

In a 2015 article published in *Science* (Open Science Collaboration, 2015), a research team undertaking a “large-scale, collaborative effort to obtain an initial estimate of the reproducibility of psychological science” (p. aac4716-1) found that although 97% of the original psychological studies in their sample had reported statistically significant results ($p < 0.05$), only 36% could be independently replicated at a statistically significant level. Some reasons for this result, according to the investigators, included procedures that were used in the MISTRA IQ study: “Selective reporting, selective analysis, and insufficient specification of the conditions necessary or sufficient to obtain the results” (p. aac4716-1).

Three years earlier, Leslie John et al. (2012) surveyed over 2,000 psychologists about their involvement in “questionable research practices” such as failing to report data, and manipulating data and statistics. A “surprisingly high” (p. 530) percentage of anonymous respondents

admitted to having engaged in such practices. John and colleagues concluded that “some questionable practices may constitute the prevailing research norm” in the field of psychology (p. 524).

In the MISTRA researchers’ defense, many investigators engaging in what we now call *p*-hacking practices may not have seen anything particularly egregious in the practice, because apparently it was part of the accepted culture in which they operated. As psychologist Joseph Simmons et al. (2018) wrote, *p*-hacking was seen by many of their colleagues as an offense similar to jaywalking:

After much discussion, our best guess was that so many published findings were false because researchers were conducting many analyses on the same data set and just reporting those that were statistically significant, a behavior that we later labeled “*p*-hacking” ... We knew many researchers – including ourselves – who readily admitted to dropping dependent variables, conditions, or participants to achieve significance. Everyone knew it was wrong, but they thought it was wrong the way it is wrong to jaywalk. (p. 255)

As its name suggests, researchers in the field of behavioral genetics expect to find genetic influences on behavior. The main task, as most see it, is to determine the correct heritability estimate for the behavioral characteristic in question. Although in 2016, Plomin and colleagues argued that the findings supposedly produced by their field of behavioral genetics “have replicated robustly” (Plomin et al., 2016, p. 3), the field may soon be approaching not a replication crisis, but an *assumption* crisis (Joseph, 2015, in press).

A Key MISTRA Assumption Is Not Supported by the Evidence

We have seen that the 1990 *Science* study 70% IQ heritability estimate was based on the idea that the MZA correlation alone “directly estimates heritability,” given the MISTRA assumption that “all resemblance between reared apart relatives is because of genetic factors” (McGue & Bouchard, 1989, p. 22). This assumption, however, overlooks the role of cohort influences as a major environmental confound in TRA studies.

Even perfectly separated MZA pairs share many *nonfamilial* environmental influences in common. The *cohort effect* concept refers to similarities in age-matched people’s IQ scores, 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 (Elder & Shanahan, 2006; Elder et al., 2015; Rose, 1982; Schaie, 1965).

Hypothetical MZA pairs separated during the first few months of life, who never met each other, and who spent their entire lives not knowing they had a twin, share many of the 15 prenatal and postnatal *nonfamilial* behavior-influencing environmental factors seen in Table 3.

Bouchard and colleagues were the first TRA investigators to recognize, as Kamin (1974) compelled them to recognize, that “for most psychological, physiological, and medical variables there are substantial age and sex effects” (McGue & Bouchard, 1984, p. 325). They devised a questionable and complicated statistical procedure, which used a “twin-based approach” to correct their correlations for these effects (p. 325). Apart from these age and sex corrections, Bouchard and colleagues assumed that none of the influences seen in Table 3 increased MZA IQ correlations for nongenetic reasons.

In order to conclude that above-zero MZA IQ test score correlations are caused only by genetic influences, TRA researchers must control for the potential environmental confounds and cohort influences seen in Table 3, or they must assume that these environmental confounds and cohort influences do not exist. For the most part, the MISTRA researchers chose the latter course.

The Environment Is Mostly Genetic

We have seen that after bypassing their model-fitting procedures and their DZA control group data (Fig. 1), Bouchard and colleagues based their conclusions about IQ heritability on the claim that the MZA correlation alone “directly estimates heritability.” However, they reached their conclusions only because they decided to count most environmental influences as *genetic* influences.

The 1990 Science Article

Bouchard and colleagues wrote in their 1990 *Science* article that one of the three “implications” of their genetic “findings,” and of behavioral genetic findings in general, was 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 (pp. 227–228). They continued,

Table 3. Fifteen nonfamilial environmental influences experienced or potentially experienced by monozygotic twin pairs separated near birth and first reunited when studied

Nonfamilial environment influence(s)	Reference(s)
1. Prenatal, including prenatal exposure to toxins and other influences	Grandjean and Landrigan (2014), Knopik et al. (2016), and Séguin et al. (2003)
2. Perinatal-infancy health care, nutrition, and exposure to environmental toxins	Lanphear (2015), Liu and Lewis (2014), and Reuben et al. (2017)
3. Birth cohort, same age, which in IQ terms might create more similar scores based on the “Flynn effect” and exposure to similar methods of education	Elder et al. (2015), Flynn (1984, 1999): “Massive IQ gains over time revealed that the present generation has a huge IQ advantage over the previous generation” (p. 5); and Schaie (1965)
4. Selective placement status (adoption)	Kamin (1974) and Richardson and Norgate (2006)
5. Gender cohort (same sex)	Maccoby and Jacklin (1975), Reilly (2012), and Zosuls et al. (2011)
6. Developmental stage, maturational change	Elder et al. (2015) and Schaie (1965)
7. Striking physical resemblance, including facial appearance and height	Cropanzano and James (1990), Hu et al. (2018), and Zebrowitz and Montepare (2008)
8. Condition of being an adopted child, with accompanying potential abandonment, attachment, and mental health issues	Keys et al. (2008) and Newman et al. (2015)
9. National, regional, ethnic, religious, and political culture	Elder et al. (2015) and Wuthnow (2007)
10. Socioeconomic status	Bradley and Corwyn (2002), Jednoróg et al. (2012), and Moriguchi and Shinohara (2019)
11. Oppression, racism, discrimination, or privilege on the basis of common racial or national background, gender, socioeconomic status, disabilities, etc.	David and Derthick (2018), Pascoe and Smart Richman (2009), and Rosenkranz and Edwin (2009)
12. Restricted range of adoptive family environments	Cropanzano and James (1990) and Stoolmiller (1999)
13. Shifting gender roles and increased career opportunities for women; access to birth control	Eccles (1987) and Welling (2013)
14. Exposure to the mass media, Internet, social media, etc.	Firth et al. (2019) and Huesmann et al. (2003)
15. Diet/nutrition	O’Neil et al. (2014) and Spencer et al. (2017)

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. (Bouchard et al., 1990a, pp. 227–228)

The above statement is not an “implication” of the researchers’ findings; rather, it is an assumption upon which they based their findings.

In the 1990 *Science* IQ study, the researchers accepted as “probably” true the “radical environmentalist” position that “the proximal cause of most psychological variance ... involves learning through experience” (p. 227). However, using an argument that sociologist Aaron Panofsky (2014) called “clever intellectual jujitsu” (p. 146), they transformed these environmental causes into genet-

ic causes based on their claim that exposure to similar environments is “guided by the steady pressure of the genome (a more distal cause)” (Bouchard et al., 1990a, p. 227).

Later Statements by Segal and Bouchard

Segal (2012) restated the MISTRA most-environmental-causes-are-really-genetic-causes position in *Born Together – Reared Apart*: “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” (p. 111). In other words, the

MISTRA researchers assumed the genetic basis of behavioral differences in the process of testing for it.

In *Born Together – Reared Apart*, Segal (2012) wrote that the MISTRA IQ study found that “shared genes, rather than social contact between family members, explained the twins’ resemblance in measured intelligence [IQ]” (p. 109). This finding, however, as Segal wrote in the same book, was based on a MISTRA *assumption* that “shared genes underlie similarity between relatives” (p. 63). The paraphrased MISTRA argument went something like this: Our findings show that shared genes explain twins’ resemblance in measured intelligence, based on our assumption that shared genes explain twins’ resemblance in measured intelligence. Indeed, Segal (2012) defended the study’s conclusions on the basis of this argument, writing that “Bouchard agreed with our critics that MZA twins probably do have more similar environments than DZA twins, *but they are environments of the twins’ making*” (p. 104, emphasis in original).

Looking back on the MISTRA in 2016, Bouchard reaffirmed that he and his colleagues “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” (Bouchard, 2016, pp. 73–74). Yet we have seen that in 1990, Bouchard calculated his 70% IQ heritability estimate “under the assumption of no environmental similarity.”

Gladys and Helen

In his 1976 chapter on intelligence (Bouchard, 1976), published three years prior to initiating the MISTRA, Bouchard mentioned Newman and colleagues’ 1937 MZA pair “Gladys and Helen.” Gladys had a Stanford-Binet IQ score of 92, whereas Helen’s score was 116, a difference of 24 points (Newman et al., 1937). Bouchard attributed the differing IQ scores between these genetically identical sisters to the differing environments experienced by each twin, with no attempt to claim that these environmental influences should be counted as genetic influences. “The twenty-four-point IQ difference,” Bouchard wrote, “can best be understood as a poor genotype-by-environment fit. Gladys’ genotype probably would have permitted the development of an IQ of 116, but it was not allowed to develop because of a less-than-favorable environment” (Bouchard, 1976, p. 175). Bouchard, in 1976, recognized that being raised in a “less-than-favorable” environment – as opposed to a favorable

one – could lower a person’s expected IQ score by 24 points. For the pre-MISTRA Bouchard, environmental influences were just that – environmental influences – and at least in this example, for him they had a powerful effect on determining a person’s IQ score.

The MISTRA researchers arrived at their conclusion that IQ is “strongly affected by genetic factors” first by suppressing their control group results, then by denying or minimizing the impact of cohort influences and other environmental factors, followed by counting most of the environmental influences they did recognize as genetic influences.

Data Kept Secret

Another issue was the Minnesota researchers’ decision to deny access to critically minded reviewers seeking to perform independent inspection and analysis of the MISTRA raw data and the twins’ case histories (Dorfman, 1995). Chambers (2017) referred to this practice as “data hoarding.” *Data sharing*, on the other hand, “makes it easier to detect questionable research practices” (Chambers, 2017, p. 77). Granting access to independent reviewers is especially important in the MISTRA case because, unlike other types of human behavioral research, TRA studies are very difficult to reproduce because reared-apart twin pairs are extremely rare. Due to changing policies, practices, and social conditions, most likely it will never again be possible for anyone to collect large enough MZA and DZA samples to conduct a new TRA study.

Bouchard once told a journalist that he “wouldn’t let Leon Kamin anywhere near” the MISTRA raw research material, although he was available to answer a “legitimate question” (Wright, 1997, p. 69). Bouchard also turned down separate requests by psychologist William Tucker and neuroscientist Steven Rose to review the MISTRA raw data (Rose, 1997; Tucker, 1997). Segal (2012) believed 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 (p. 299). Critics were placed 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 (Joseph, 2015). At the same time, Bouchard and Segal granted access to the MISTRA raw data to friends and “trusted” colleagues (Segal, 2012, p. 184).

Because critically minded scholars were not allowed to inspect and analyze the MISTRA raw data, including twins’ unpublished case histories and information on

their degree of separation, by default and until proven otherwise we must assume that the MISTRA MZA pairs were no more “reared apart” than were what Farber called the *partially* reared-apart MZA samples found in the original three TRA studies.

Regardless of their reasons for denying access, when TRA researchers fail to make their unpublished raw data and the twins’ life history information available for inspection and analysis by qualified independent reviewers, we must evaluate their findings with extreme caution – or even reject their findings outright – because these findings are based on virtually unreproducible sample populations. Transparency is even more imperative when researchers’ conclusions have important educational, social relations, and political policy implications, and when they are cited and promoted by far-right white-nationalist political groups (e.g., Taylor, 1993) in support of these groups’ agendas and actions (Lerner & Chase, 2020). IQ-hereditarian research is a frequent topic of online discussions among supporters of such groups (Panofsky et al., 2021). The MISTRA itself was largely financed by the Pioneer Fund (about 60% or USD 1.42 million, roughly USD 3.5 million in 2022 US dollars; Bouchard et al., 1990a; Segal, 2012), an organization with a history of promoting racial differences research and eugenics (Joseph, 2015; Lombardo, 2002; Tucker, 2002).

More on Genetic Confirmation Bias

In a footnote in the 2007 Johnson et al. MISTRA publication on mental abilities, the researchers revealed that their software program was designed to diminish the statistical weight given to unexpectedly high DZA group test score correlations. They could have concluded that no genetic influences were detected, but they instead chose to assume that these high DZA correlations were the result of “sampling variability.” Clearly referring to their MZA and DZA pairs, they wrote,

For four 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. (Johnson et al., 2007, p. 551, footnote 4)

It appears that genetic confirmation bias was built into the MISTRA computer software program (Joseph, 2015). Once again, the investigators assumed the operation of genetic influences in the process of testing for it. When researchers interpret and define most of what they see as

being caused by genetics – including what they recognize as “proximal” environmental influences – conclusions in favor of genetic causes and moderate-to-high heritability estimates necessarily will follow.

Bouchard and McGue (2003) stated that “behavior genetic methods are unbiased with regard to whether genetic or environmental sources of variance are more important” (p. 5). And according to Segal (2012), “the MISTRA investigators were often accused of looking for genetic factors, emphasizing the twins’ similarities and ignoring their differences. This was not the case – 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*” (p. 244, emphasis in original). However, “the twins” did not decide (a) to call partially reared-apart twins “reared-apart twins,” (b) to base the study’s findings on key assumptions “likely not to hold,” (c) to suppress the study’s control group IQ correlations, (d) to conclude that a significantly above-zero MZA behavioral correlation can only be caused by genetic factors based on an assumption stating the very same thing, (e) to largely ignore the confounding influence of cohort effects and other nongenetic factors, (f) to count most environmental influences as genetic influences, and (g) to deny independent scholars access to the MISTRA raw data – the researchers did.

Summary and Conclusions

Apart from the controversies surrounding concepts such as IQ, general intelligence, heritability, and model fitting, and the questionable status of the MISTRA pairs as “reared-apart” twins as well as other TRA study problems discussed earlier, I have attempted to show that the 1990 *Science* IQ study conclusion that “general intelligence or IQ is strongly affected by genetic factors” (70% heritability) was based on several of the researchers’ key decisions.

1. As seen in Figure 1, they decided to omit and bypass their step 2 DZA control group IQ score correlations and to base their conclusions on the unsupported step 3B assumption that heritability “can be directly estimated from the MZA correlation” alone (Bouchard et al., 1990a, p. 224).
2. Also seen in Figure 1, they decided to bypass their step 3A model-fitting procedure (which used MZA, DZA, and other data), even though they used this procedure in most non-IQ MISTRA publications appearing prior to the 1990 IQ study (e.g., Tellegen et al., 1988), after

the IQ study (e.g., DiLalla et al., 1996), and continuing into the 21st century (e.g., Koenig et al., 2005).

3. They decided to assume that numerous potential *non-familial* behavior-shaping influences such as prenatal, adoptee-placement, cultural, religious, similar physical appearance, and other cohort influences shared by MZA pairs either do not exist, or that most should be counted as genetic influences. (The researchers' age and sex effect correlation corrections were their only attempt to address cohort influences.)

I have suggested that even though Bouchard and colleagues began their 1990 *Science* IQ study with a statement that "MZA and DZA twin pairs ... provide the simplest and most powerful method for disentangling the influence of environmental and genetic factors on human characteristics," had they chosen to use this "most powerful method" they would have had to conclude that they found no evidence in favor of IQ heritability. Segal (2012) described a TRA study process in which intraclass "correlations are calculated separately for MZA and DZA twin pairs and compared" (p. 59). Yet in the MISTRA IQ study, MZA and DZA twin correlations *were not* compared.

In 1980, sociologist Howard Taylor described what he called "The IQ game," by which he meant IQ-genetic researchers' "use of assumptions that are implausible as well as arbitrary to arrive at some numerical value for the genetic heritability of human IQ scores on the grounds that no heritability calculations could be made without benefit of such assumptions" (Taylor, 1980, p. 7). The MISTRA IQ study can be seen as an exemplar of "IQ game" bad science.

Some people might defend the MISTRA IQ study's conclusions on the grounds that, as mentioned in the *Science* article abstract, researchers performing other types of behavioral genetic studies arrived at similar conclusions. There are at least two ways to counter such a point. (a) A psychological study, test, or method must stand or fall on its own logic and soundness, and cannot be validated by supposed "converging evidence" from other methods (Lilienfeld et al., 2003). (b) Several authors have challenged the findings of the other TRA studies (e.g., Farber, 1981; Joseph, 2001, 2015; Kamin, 1974; Kamin & Goldberger, 2002; Lerner, 2018; Lewontin et al., 1984; Taylor, 1980), "twin method" MZT-DZT comparisons (e.g., Charney, 2012; Joseph, 2015; Richardson & Norgate, 2005), (nontwin) adoption studies (e.g., Joseph, 2004, 2006; Kamin, 1974; Richardson & Norgate, 2006; Rutter, 2006; Stoolmiller, 1999), and molecular genetic "genome-wide association" and "polygenic risk score" studies (e.g., Baverstock, 2019; Richardson & Jones, 2019).

Although the MISTRA has thus far avoided serious mainstream academic scrutiny, such scrutiny is long overdue. Ritchie (2020) described psychology's "badly broken" (p. 6) research/publication system and discussed studies that "involved a scientific 'finding' that, upon closer scrutiny, turned out to be either less solid than it seemed, or to be completely untrue" (p. 34). Although Ritchie excluded behavioral genetic twin and adoption studies from his analysis, upon closer examination, the MISTRA IQ study failed to discover evidence that genetic factors influence IQ scores and cognitive ability across the studied population.

Acknowledgements

I thank Richard M. Lerner and an anonymous reviewer for their very helpful comments on the original draft of this paper. R.M.L. provided additional help as I was preparing the revision. All views expressed in this article are entirely my own, and any errors are my responsibility.

Statement of Ethics

No human or animal participants were used in this study. There was no study protocol to approve.

Conflict of Interest Statement

The author has no conflicts of interest.

Funding Sources

This study was not funded by anyone other than the author.

Author Contributions

Solo-authored work. Some parts of the submitted work have been adapted from the author's non-peer-reviewed online articles. Areas adapted from the author's 2015 book on twin research are cited (Joseph, 2015).

Data Availability Statement

All data were taken from previously published sources and re-analyzed by the author.

References

- Angle, R., & Neimark, J. (1997, July). Nature's clones: Is the gene king? Studies of twins find that genes shape multiple areas including personality. *Psychology Today*. <https://www.psychologytoday.com/us/articles/199707/natures-clone>
- Arvey, R. D., Bouchard, T. J., Jr., Segal, N. L., & Abraham, L. M. (1989). Job satisfaction: Environmental and genetic components. *Journal of Applied Psychology, 74*, 187–192.
- Baker, M. (2015, August 27). Over half of psychology studies fail reproducibility test. *Nature*.
- Baverstock, K. (2019). Polygenic scores: Are they a public health hazard? *Progress in Biophysics and Molecular Biology, 149*, 4–8.
- Begley, S. (1987, November 23). *All about twins*. Newsweek, pp. 58–69.
- Bernstein, D. A., Clarke-Stewart, A., Penner, L. A., Roy, E. J., & Wickens, C. D. (2000). *Psychology* (5th ed.). Houghton Mifflin.
- Billings, P. R., Beckwith, J., & Alper, J. S. (1992). The genetic analysis of human behavior: A new era? *Social Science and Medicine, 35*, 227–238.
- Bouchard, T. J., Jr. (1976). Genetic factors in intelligence. In A. R. Kaplan (Ed.), *Human behavior genetics* (pp. 164–197). Thomas.
- Bouchard, T. J., Jr. (1981). The study of mental ability using twin and adoption designs. In L. Gedda, P. Parisi, & W. E. Nance (Eds.), *Twin research 3: Part B. Intelligence, personality, and development* (pp. 21–23). Liss.
- Bouchard, T. J., Jr. (1984). Twins reared together and apart: What they tell us about human diversity. In S. W. Fox (Ed.), *Individuality and determinism: Chemical and biological bases* (pp. 147–184). Plenum Press.
- Bouchard, T. J., Jr. (1993). The genetic architecture of human intelligence. In P. A. Vernon (Ed.), *Biological approaches to the study of human intelligence* (pp. 33–93). Ablex Publishing Corporation.
- 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*, 415–421.
- 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). Cambridge University Press.
- Bouchard, T. J., Jr. (1998). Genetic and environmental influences on adult intelligence and special mental abilities. *Human Biology, 70*, 257–279. <https://www.jstor.org/stable/41465638>
- Bouchard, T. J., Jr. (2016). Genes and behavior: Nature via nurture. In R. J. Sternberg, S. T. Fiske, & D. J. Foss (Eds.), *Scientists making a difference: One hundred eminent behavior and brain scientists talk about their most important contributions* (pp. 73–76). Cambridge University Press.
- Bouchard, T. J., Jr., Lykken, D. T., McGue, M., Segal, N. L., & Tellegen, A. (1990a). Sources of human psychological differences: The Minnesota Study of Twins Reared Apart. *Science, 250*, 223–228.
- Bouchard, T. J., Jr., Lykken, D. T., Segal, N. L., & Wilcox, K. J. (1986). Development in twins reared apart: A test of the chronogenetic hypothesis. In A. Demirjian & M. Brault Dubuc (Eds.), *Human growth: A multidisciplinary review* (pp. 299–310). Taylor & Francis.
- Bouchard, T. J., Jr., & McGue, M. (1981). Familial studies of intelligence: A review. *Science, 212*, 1055–1059.
- 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.
- Bouchard, T. J., Jr., & McGue, M. (2003). Genetic and environmental influences on human psychological differences. *Journal of Neurobiology, 54*, 4–45.
- Bouchard, T. J., Jr., Segal, N. L., & Lykken, D. T. (1990b). Genetic and environmental influences on special mental abilities in a sample of twins reared apart. *Acta Geneticae Medicae et Gemellologiae (Twin Research), 39*, 193–206.
- Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53*, 371–399.
- Brody, J. (2018, August 20). What twins can teach us about nature vs. nurture. New York Times. <https://www.nytimes.com/2018/08/20/well/family/what-twins-can-teach-us-about-nature-vs-nrture.html>
- Carey, B. (2011, November 2). Fraud case seen as a red flag for psychology research. New York Times. <https://www.nytimes.com/2011/11/03/health/research/noted-dutch-psychologist-stapel-accused-of-research-fraud.html>
- Chambers, C. (2017). *The seven deadly sins of psychology: A manifesto for reforming the culture of scientific practice*. Princeton University Press.
- Charney, E. (2012). Behavior genetics and postgenomics. *Behavioral and Brain Sciences, 35*, 331–358.
- Cropanzano, R., & James, K. (1990). Some methodological considerations for the behavioral genetic analysis of work attitudes. *Journal of Applied Psychology, 75*, 433–439.
- David, E. J. R., & Derthick, A. O. (2018). *The psychology of oppression*. Springer.
- deBoer, F. (2020). *The cult of smart: How our broken system perpetuates social injustice*. All Points Books.
- DiLalla, D. L., Gottesman, I. I., Carey, G., & Bouchard, T. J., Jr. (1996). Heritability of MMPI personality indicators of psychopathology in twins reared apart. *Journal of Abnormal Psychology, 105*, 491–499.
- Dorfman, D. D. (1995). Soft science with a neo-conservative agenda [Review of the book *The bell curve: Intelligence and class structure in American life*, by R. J. Herrnstein & C. Murray]. *Contemporary Psychology, 40*, 418–421.
- Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychology of Women Quarterly, 11*, 135–172.
- Elder, G. H., Jr., & Shanahan, M. J. (2006). The life course and human development. In R. M. Lerner (Ed.), *Theoretical models of human development. Handbook of child psychology* (6th ed., Vol. 1, pp. 665–715). Wiley.
- Elder, G. H., Jr., Shanahan, M. J., & Jennings, J. A. (2015). Human development in time and place. In M. H. Bornstein & T. Leventhal (Eds.), *Handbook of child psychology and developmental science. Ecological settings and processes in developmental systems* (7th ed., Vol. 4, pp. 6–54). Wiley.
- Fancher, R. E. (1985). *The intelligence men: Makers of the IQ controversy*. Norton.
- Farber, S. L. (1981). *Identical twins reared apart: A reanalysis*. Basic Books.
- Firth, J., Torous, J., Stubbs, B., Firth, J. A., Steiner, G. Z., Smith, L., Alvarez-Jimenez, M., Gleeson, J., Vancampfort, D., Armitage, C. J., & Sarris, J. (2019). The “online brain”: How the Internet may be changing our cognition. *World Psychiatry, 18*(2), 119–129.
- Fischer, C. S., Hout, M., Sánchez Jankowski, M., Lucas, S. R., Swidler, A., & Voss, K. (1996). *Inequality by design: Cracking the bell curve myth*. Princeton University Press.
- Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95*, 29–51.
- Flynn, J. R. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist, 54*, 5–20.
- Ford, B. D. (1993). Emergenesis: An alternative and a confound [Letter to the editor]. *American Psychologist, 48*, 1294.
- Gold medal award for life achievement in the science of psychology (2014). *American Psychologist, 69*, 477–479.
- Goldberger, A. S. (1979). Heritability. *Economica, 46*, 327–347.
- Gould, S. J. (1981). *The mismeasure of man*. Norton.
- Grandjean, P., & Landrigan, P. J. (2014). Neurobehavioural effects of developmental toxicity. *The Lancet Neurology, 13*(3), 330–338.
- Harris, J. R. (1998). *The nurture assumption: Why children turn out the way they do*. The Free Press.
- Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of p-hacking in science. *PLoS Biology, 13*(3), e1002106.
- Hearnshaw, L. S. (1979). *Cyril Burt: Psychologist*. Cornell University Press.
- Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. The Free Press.
- Hewstone, M., Fincham, F. D., & Foster, J. (2005). *Psychology*. Blackwell.
- Holden, C., & Bouchard, T. (2009). Newsmaker interview: Behavioral geneticist celebrates twins, scorns PC science. *Science, 325*(5936), 27.

- Hu, Y., Parde, C. J., Hill, M. Q., Mahmood, N., & O'Toole, A. J. (2018). First impressions of personality traits from body shapes. *Psychological Science, 29*, 1969–1983.
- Huesmann, L. R., Moise-Titus, J., Podolski, C., & Eron, L. D. (2003). Longitudinal relations between children's exposure to TV violence and their aggressive and violent behavior in young adulthood: 1977–1992. *Developmental Psychology, 39*, 201–221.
- Hur, Y. M., & Bouchard, T. J., Jr. (1995). Genetic influences on perceptions of childhood family environment: A reared apart twin study. *Child Development, 66*, 330–345.
- Jednoróg, K., Altarelli, I., Monzalvo, K., Fluss, J., Dubois, J., Billard, C., Dehaene-Lambertz, G., & Ramus, F. (2012). The influence of socioeconomic status on children's brain structure. *PLoS One, 7*(8), e42486.
- Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review, 39*, 1–123.
- Jinks, J. L., & Fulker, D. W. (1970). Comparison of the biometrical genetic, MAVA, and classical approaches to the analysis of human behavior. *Psychological Bulletin, 73*, 311–349.
- John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science, 23*, 524–532.
- Johnson, W., Bouchard, T. J., Jr., McGue, M., Segal, N. L., Tellegen, A., Keyes, M., & Gottesman, I. I. (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.
- Jonsson, H., Magnusdottir, E., Eggertsson, H. P., Stefansson, O. A., Arnadottir, G. A., Eiriksson, O., Zink, F., Helgason, E. A., Jonsdottir, I., Gylfason, A., Jonasdottir, A., Jonasdottir, A., Beyter, D., Steingrimsdottir, T., Norddahl, G. L., Magnusson, O. T., Masson, G., Halldorsson, B. V., Thorsteinsdottir, U., & Stefansson, K. (2021). Differences between germline genomes of monozygotic twins. *Nature Genetics, 53*, 27–34.
- Joseph, J. (2001). Separated twins and the genetics of personality differences: A critique. *American Journal of Psychology, 114*, 1–30.
- Joseph, J. (2004). *The gene illusion: Genetic research in psychiatry and psychology under the microscope*. Algora.
- Joseph, J. (2006). *The missing gene: Psychiatry, heredity, and the fruitless search for genes*. Algora.
- Joseph, J. (2015). *The trouble with twin studies: A reassessment of twin research in the social and behavioral sciences*. Routledge.
- Joseph, J. (2018, April 4). *Leon J. Kamin (1927–2017): A nemesis of genetic determinism and scientific racism*. Mad in America. <https://www.madinamerica.com/2018/04/leon-j-kamin-nemesis-genetic-determinism/>
- Joseph, J. (in press). A “blueprint” for genetic determinism: An appraisal of Robert Plomin's *Blueprint: How DNA Makes Us Who We Are* [Review of the book *Blueprint: How DNA makes us who we are*, by R. Plomin]. *American Journal of Psychology*. <https://cpb-us-e1.wpmucdn.com/sites.ucsc.edu/dist/0/158/files/2021/10/FinalJay-Joseph-Review-of-Blueprint-Manuscript-AJP-Revised-9-17-2021.pdf>
- Joseph, J., & Baldwin, S. (2000). Four editorial proposals to improve social sciences research and publication. *International Journal of Risk and Safety in Medicine, 13*, 109–116.
- Juel-Nielsen, N. (1965/1980). *Individual and environment: Monozygotic twins reared apart* (revised ed.). International Universities Press.
- Kamin, L. J. (1974). *The science and politics of IQ*. Erlbaum.
- Kamin, L. J., & Goldberger, A. S. (2002). Twin studies in behavioral research: A skeptical view. *Theoretical Population Biology, 61*, 83–95.
- Keller, E. F. (2010). *The mirage of a space between nature and nurture*. Duke University Press.
- Keller, L. M., Arvey, R. D., Bouchard, T. J., Jr., Segal, N. L., & Dawis, R. V. (1992). Work values: Genetic and environmental influences. *Journal of Applied Psychology, 77*, 79–88.
- Keys, M. A., Sharma, A., Elkins, I. J., Iacono, W. G., & McGue, M. (2008). The mental health of US adolescents adopted in infancy. *Archives of Pediatric & Adolescent Medicine, 162*, 419–425.
- Knopik, V. S., Neiderhiser, J. M., de Geus, E., & Boomsma, D. (2016). The importance of the prenatal environment in behavioral genetics: Introduction to special issue. *Behavior Genetics, 46*, 281–285.
- Koenig, L. B., McGue, M., Krueger, R. F., & Bouchard, T. J., Jr. (2005). Genetic and environmental influences on religiosity: Findings for retrospective and current religiosity ratings. *Journal of Personality, 73*, 471–488.
- Lanphear, B. P. (2015). The impact of toxins on the developing brain. *Annual Review of Public Health, 36*, 211–230.
- Lerner, R. M. (2018). *Concepts and theories of human development* (4th ed.). Routledge.
- Lerner, R. M., & Chase, P. A. (2020). Hate in contemporary America: Pathology or opportunism? In R. J. Sternberg (Ed.), *Perspectives on hate: How it originates, develops, manifests, and spreads* (pp. 137–160). American Psychological Association.
- Lewis, T. (2019, August 14). *Twin brothers separated at birth reveal striking genetic similarities*. CBS News. <https://www.cbsnews.com/news/twin-brothers-separated-at-birth-reveal-striking-genetic-similarities/>
- Lewontin, R. C. (1974). The analysis of variance and the analysis of causes. *American Journal of Human Genetics, 26*, 400–411.
- Lewontin, R. C., Rose, S., & Kamin, L. J. (1984). *Not in our genes*. Pantheon.
- Lilienfeld, S. O., Lynn, S. J., & Lohr, J. M. (2003). Science and pseudoscience in clinical psychology: Initial thoughts, reflections, and considerations. In S. O. Lilienfeld, S. J. Lynn, & J. M. Lohr (Eds.), *Science and pseudoscience in clinical psychology* (pp. 1–14). Guilford.
- Liu, J., & Lewis, G. (2014). Environmental toxicity and poor cognitive outcomes in children and adults. *Journal of Environmental Health, 76*(6), 130–138.
- Lombardo, P. A. (2002). “The American breed”: Nazi eugenics and the origins of the Pioneer Fund. *Albany Law Review, 65*, 743–830.
- Lykken, D. T. (1995). *The antisocial personalities*. Erlbaum.
- Maccoby, E. E., & Jacklin, C. N. (1975). *The psychology of sex differences*. Stanford University Press.
- McCourt, K., Bouchard, T. J., Jr., Lykken, D. T., Tellegen, A., & Keyes, M. (1999). Authoritarianism revisited: Genetic and environmental influences in twins reared apart and together. *Personality and Individual Differences, 27*, 985–1014.
- McGue, M., & Bouchard, T. J., Jr. (1984). Adjustment of twin data for the effects of age and sex. *Behavior Genetics, 14*, 325–343.
- McGue, M., & Bouchard, T. J., Jr. (1989). Genetic and environmental determinants of information processing and special mental abilities: A twin analysis. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 5, pp. 7–45). Erlbaum.
- Mensh, E., & Mensh, H. (1991). *The IQ mythology: Class, race, gender, and inequality*. Southern Illinois Press.
- Moore, D. S., & Shenk, D. (2016). The heritability fallacy. *WIREs Cognitive Science*.
- Moriguchi, Y., & Shinohara, I. (2019). Socioeconomic disparity in prefrontal development during early childhood. *Scientific Reports, 9*, 2585.
- Mukherjee, S. (2016). *The gene: An intimate history*. Scribner.
- Myers, D. G. (2005). *Exploring psychology* (6th ed., in modules). Worth Publishers.
- 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.
- Newman, H. H., Freeman, F. N., & Holzinger, K. J. (1937). *Twins: A study of heredity and environment*. University of Chicago Press.
- Newman, L., Sivaratnam, C., & Komiti, A. (2015). Attachment and early brain development – neuroprotective interventions in infant-caregiver therapy. *Translational Developmental Psychiatry, 3*, 1–12.
- O'Grady, C. (2020, July 15). *Misconduct allegations push psychology hero off his pedestal*. Science. <https://www.sciencemag.org/news/2020/07/misconduct-allegations-push-psychology-hero-his-pedestal>

- O'Neil, A., Quirk, S. E., Housden, S., Brennan, L., Williams, L. J., Pasco, J. A., Berk, M., & Jacka, F. N. (2014). Relationship between diet and mental health in children and adolescents: A systematic review. *American Journal of Public Health, 104*(10), e31–e42.
- Open Science Collaboration (2015). Psychology: Estimating the reproducibility of psychological science. *Science, 349*(6251), aac47161–aac47168.
- Panofsky, A. (2014). *Misbehaving science: Controversy and the development of behavior genetics*. University of Chicago Press.
- Panofsky, A., Dasgupta, K., & Iturriaga, N. (2021). How White nationalists mobilize genetics: From genetic ancestry and human biodiversity to counterscience and metapolitics. *American Journal of Physical Anthropology, 175*, 387–398.
- Pascoe, E. A., & Smart Richman, L. (2009). Perceived discrimination and health: A meta-analytic review. *Psychological Bulletin, 135*, 531–554.
- Pedersen, N. L., Plomin, R., Nesselroade, J. R., & McClearn, G. E. (1992). A quantitative genetic analysis of cognitive abilities during the second half of the life span. *Psychological Science, 3*, 346–353.
- Pinker, S. (2002). *The blank slate*. Viking.
- Plomin, R. (2004). *Nature and nurture: An introduction to human behavioral genetics*. Thompson Wadsworth.
- Plomin, R., Chipuer, H. M., & Loehlin, J. C. (1990). Behavioral genetics and personality. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 225–243). Guilford.
- Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2013). *Behavioral genetics* (6th ed.). Worth Publishers.
- Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2016). Top 10 replicated findings from behavioral genetics. *Perspectives on Psychological Science, 11*, 3–23.
- Purcell, S. (2013). Statistical methods in behavioral genetics. In R. Plomin, J. C. DeFries, V. S. Knopik, & J. M. Neiderhiser (Eds.), *Behavioral genetics* (6th ed., pp. 357–411). Worth Publishers.
- Raven, J. C. (1986). *Manual for raven's progressive matrices and vocabulary scales*. Lewis.
- Reilly, D. (2012). Gender, culture, and sex-typed cognitive abilities. *PLoS One, 7*(7), e39904.
- Reuben, A., Caspi, A., Belsky, D. W., Broadbent, J., Harrington, H., Sugden, K., Houts, R. M., Ramrakha, S., Poulton, R., & Moffitt, T. E. (2017). Association of childhood blood lead levels with cognitive function and socioeconomic status at age 38 years and with IQ change and socioeconomic mobility between childhood and adulthood. *JAMA, 317*(12), 1244–1251.
- Richardson, K. (2017). *Genes, brains, and human potential*. Columbia University Press.
- Richardson, K., & Norgate, S. (2005). The equal environments assumption of classical twin studies may not hold. *British Journal of Educational Psychology, 75*, 339–350.
- Richardson, K., & Norgate, S. (2006). A critical analysis of IQ studies of adopted children. *Human Development, 49*, 319–335.
- Richardson, K., & Jones, M. C. (2019). Why genome-wide associations with cognitive ability measures are probably spurious. *New Ideas in Psychology, 55*, 35–41.
- Ridley, M. (2004). The agile gene: How nature turns on nurture. *Nature via nurture*. Perennial.
- Ritchie, S. (2020). *Science fictions: How fraud, bias, negligence, and hype undermine the search for truth*. Metropolitan Books.
- Rose, R. J. (1982). Separated twins: Data and their limits. *Science, 215*, 959–960. [Review of the book *Identical twins reared apart: A reanalysis*, by S. Farber].
- Rose, S. (1997). *Lifelines: Life beyond the genes*. Oxford University Press.
- Rosenkranz, M., & Edwin, M. (2009). *White male privilege: A study of racism in America 50 years after the Voting Rights Act* (3rd ed.). Law Dogs Books.
- Rutter, M. (2006). *Genes and behavior: Nature-nurture interplay explained*. Blackwell.
- Sapolsky, R. M. (2017). *Behave: The biology of humans at our best and worst*. Penguin.
- Schaie, K. W. (1965). A general model for the study of developmental problems. *Psychological Bulletin, 64*, 92–107.
- Segal, N. L. (1999). *Entwined lives: Twins and what they tell us about human behavior*. Dutton.
- Segal, N. L. (2012). *Born together – Reared apart: The landmark Minnesota twin study*. Harvard University Press.
- Séguin, L., Xu, Q., Potvin, L., Zunzunegui, M. V., & Frohlich, K. L. (2003). Effects of low income on infant health. *Canadian Medical Association Journal, 168*, 1533–1538.
- Shields, J. (1962). *Monozygotic twins brought up apart and brought up together*. Oxford University Press.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2018). False-positive citations. *Perspectives on Psychological Science, 13*, 255–259.
- Spencer, S. J., Korosi, A., Layé, S., Shukitt-Hale, B., & Barrientos, R. M. (2017). Food for thought: How nutrition impacts cognition and emotion. *NPJ: Science of Food, Article 1*(7), 1–8.
- 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.
- Taylor, H. F. (1980). *The IQ game: A methodological inquiry into the heredity-environment controversy*. Rutgers University Press.
- Taylor, J. (1993, August). *Genetics, personality, and race*. American Renaissance. https://www.amren.com/news/2011/07/genetics_person/
- Tellegen, A., Lykken, D. T., Bouchard, T. J., Jr., Wilcox, K. J., Segal, N. L., & Rich, S. (1988). Personality similarity in twins reared apart and together. *Journal of Personality and Social Psychology, 54*, 1031–1039.
- Tucker, W. H. (1997). Re-reconsidering Burt: Beyond a reasonable doubt. *Journal of the History of the Behavioral Sciences, 33*, 145–162.
- Tucker, W. H. (2002). *The funding of scientific racism: Wickliffe Draper and the Pioneer Fund*. University of Illinois Press.
- 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.
- Waller, N. G., Kojetin, B. A., Bouchard, T. J., Jr., Lykken, D. T., & Tellegen, A. (1990). Genetic and environmental influences on religious interests, attitudes, and values: A study of twins reared apart and together. *Psychological Science, 1*, 138–142.
- Wechsler, D. (1955). *Manual for the Wechsler adult intelligence scale*. The Psychology Corporation.
- Welling, L. L. (2013). Psychobehavioral effects of hormonal contraceptive use. *Evolutionary Psychology, 11*, 718–742.
- Wright, L. (1997). *Twins: And what they tell us about who we are*. Wiley & Sons.
- Wuthnow, R. (2007). Cognition and religion. *Sociology of Religion, 68*, 341–360.
- Wyatt, W. J. (1993). Identical twins, emergence and environments [Letter to the editor]. *American Psychologist, 48*, 1294–1295.
- Zebrowitz, L. A., & Montepare, J. M. (2008). Social psychological face perception: Why appearance matters. *Social and Personality Psychology Compass, 2*, 1497–1517.
- Zosuls, K. M., Miller, C. F., Ruble, D. N., Martin, C. L., & Fabes, R. A. (2011). Gender development research in sex roles: Historical trends and future directions. *Sex Roles, 64*, 826–842.