Classical and Molecular Genetic Research on General Cognitive Ability

BY MATT MCGUE and IRVING I. GOTTESMAN

Since the time of Charles Spearman at the beginning of the twentieth century,[1] it has been widely recognized that performance on multiple tests of cognitive ability (for example, mathematics, vocabulary, spatial relations) is positively intercorrelated, a phenomenon known as "positive manifold." While specific cognitive abilities (in mathematics or vocabulary, for example) are undeniably important, positive manifold implies the existence of a common core ability that, because it underlies multiple specific abilities, is likely to have broad implications for social and educational functioning.[2] This common core has been variously called "general intelligence," *"g,"* or—our preferred term—"general cognitive ability." In this essay, we review behavioral genetics research on GCA. Before describing that research, however, we briefly discuss why GCA is considered a fundamentally important dimension of behavior on which humans differ. We then summarize behavioral genetics research that has sought to identify and quantify the total contributions of genetic and environmental factors to individual differences in GCA as well as molecular genetics research that has sought to identify genetic variants that underlie inherited effects.

The Importance of General Cognitive Ability

Arguably, no psychological variable has received more attention from behavioral geneticists than GCA, and for good reason. GCA has a rich correlational network, implying that it may play an important role in multiple domains of functioning. The original measures of GCA were developed at the beginning of the last century in France by Alfred Binet and Theodore Simon to predict differential academic progress in school-aged children. It is consequently not surprising that GCA continues to be highly correlated with various indicators of educational attainment.[3] Yet the predictive utility of GCA is not limited to academic achievement; it is also correlated with work performance, navigating the complexities of everyday life, the absence of various social pathologies (such as criminal convictions), and even health and mortality.[4] Although the causal basis for these associations is not always known, it is nonetheless the case that research on GCA has the potential to provide insights into the origins of a wide range of important social outcomes.

Research on GCA may also help with understanding the origins of at least some forms of intellectual disability. A diagnosis of intellectual disability requires evidence of low cognitive functioning, usually an IQ test score of less than 70. Yet the 2.5 percent of individuals with IQs in this range represent a heterogeneous set of etiologies (see figure 1). A two-group model of intellectual disability was first introduced by J. A. Fraser Roberts but subsequently popularized by Edward Zigler,[5] the architect of the Head Start program. In this model, there are two major types of intellectual disability. First, a minority of individuals with an intellectual disability have suffered some significant and specific neurological trauma (of either genetic or environmental origin). Such trauma usually results in a moderate to severe intellectual deficit, with IQs of less than 50 and a distribution that is discontinuous with the normal distribution of IQ. Second, a majority of individuals with an intellectual disability have a relatively minor intellectual deficit, with IQs in the 50 to 70 range. The IQs of this latter group are continuous with the normal distribution of IQ, making it difficult in most cases to identify a specific etiology (in other words, their deficit

Matt McCue and Irving I. Gottesman, "Classical and Molecular Genetic Research on General Cognitive Ability," *The Genetics of Intelligence: Ethics and the Conduct of Trustworthy Research,* special report, *Hastings Center Report* 45, no. 5 (2015): S25-S31. DOI: 10.1002/hast.495

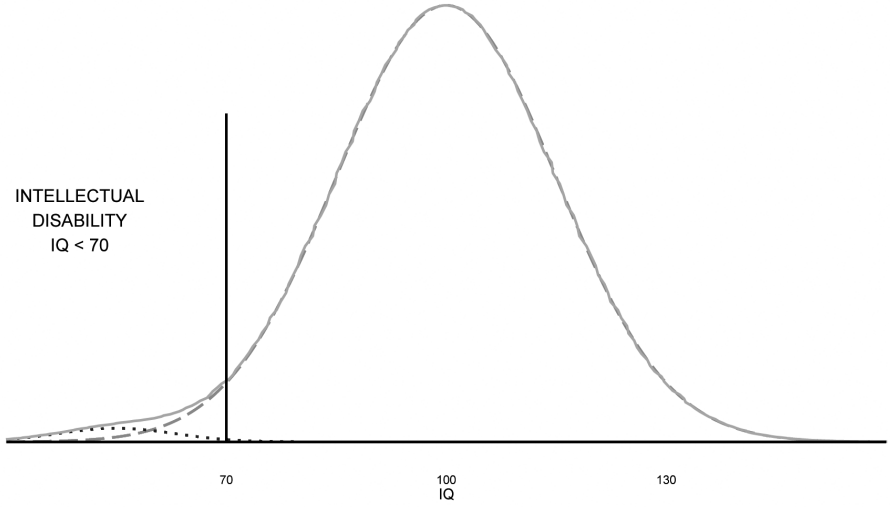


Figure 1.
Two-group model of intellectual disability

The actual distribution of IQ (depicted as a solid line) deviates slightly from a normal distribution in that there is an excess of individuals with IQs below 70, in the intellectual disability range. This excess is thought to represent two broad and distinct etiologies of intellectual disability. A minority (about 0.5 percent, depicted as a dotted line) of intellectual disability is due to some major neurological trauma, which is typically associated with the more severe levels of intellectual deficit and is discontinuous with the overall distribution of IQ. A majority of intellectual disability is due to the cumulative effect of multiple genetic and environmental factors and so is typically associated with mild levels of intellectual deficit that are continuous with the overall distribution of IQ (depicted as a dashed line).

is idiopathic) because low IQ is a result of the cumulative influence of the multiple genetic and environmental factors that also contribute to individual differences throughout the full range of IQ. One reason researchers seek to identify the genes underlying GCA is that variants in these genes may play a role with the large number of individuals with mild and idiopathic forms of intellectual disability, and knowledge of that role could eventually contribute to helping such individuals.

Classical Twin and Family Studies of General Cognitive Ability

GCA is one of the most widely studied phenotypes in behavioral genetics. More than thirty years ago, Thomas Bouchard and Matt McGue reviewed the existing literature on familial resemblance for GCA, identifying more than 500 familial correlations from more than one hundred studies on a combined sample size of greater than 100,000 familial pairings.[6] Although there have been numerous twin and family studies published since, subsequent research has not materially altered the pattern of familial resemblance for GCA from that reported in 1981.[7]

Figure 2 summarizes the weighted average GCA correlations for various familial pairings. The average correlation for reared-together, genetically identical monozygotic (MZ) twins is .86 (which approaches the test-retest correlation for the same individuals after one month). Whatever the causes of individual differences in GCA, they are clearly shared by MZ twins. The figure also shows that the observed or phenotypic correlation is related to degree of genetic relationship: as the closeness in the relationship between relatives who have been raised together increases, so does the correlation between their "levels" of GCA. For relatives that shared the same rearing or family environment, the different (and higher) observed correlations between more closely genetically related individuals suggests that genetics can help to explain what we observe. For example, if we assume that MZ and dizygotic (DZ) twins have equally similar family environments, and if we remember that MZ twins share 100 percent of their DNA, whereas DZ twins share (on average) only 50 percent of theirs, we can infer that the reason MZ twins are more similar with respect to GCA than DZ twins has something to do with their greater genetic similarity.

It is important to also recognize that DZ twins are more similar with respect to GCA than are full biological siblings, who also share, on average, 50 percent of their DNA. In turn, full biological siblings are more similar than are parent-offspring pairs, who also share 50 percent of their DNA. So, these classical twin, family, and adoption studies suggest that something more than genetics—say, shared rearing environment or developmental stage—helps to explain what we observe.

Behavioral geneticists conceptualize the environment as having two components: the *nonshared* environment and the *shared* environment. The nonshared environment refers to those aspects of the environment that contribute to differences among children in the same home with respect to an observable trait like GCA. That we need such an explanatory variable is plain when we notice that, while MZ twins raised in the same environment are highly correlated with regard to GCA (.86), they are not perfectly or 100 percent correlated (1.0). The nonshared environment is the name

General cognitive ability is correlated not only with academic achievement but also with navigating the complexities of everyday life, the absence of various social pathologies, and even health and mortality.

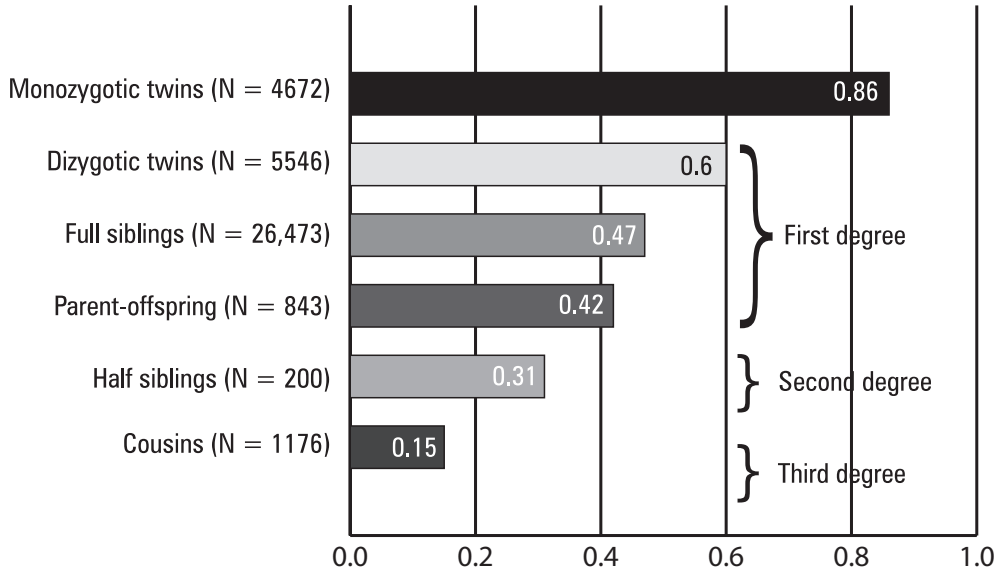
for that variable that can help to explain why individuals with the same genome and same rearing environment are not always the same with respect to some trait.

The shared environment comprises those aspects of the environment that contribute to the similarity between children in the same family with regard to a trait. Behavioral geneticists have shown the influence of the shared environment for GCA in several ways. For example, genetically related individuals who are reared together are more similar with regard to GCA than individuals who are equally genetically related but reared in separate homes (see figure 3); something is going on that makes individuals who share a home environment more similar than those who do not. Behavioral geneticists have also implicated the importance of the shared environment by showing that genetically unrelated siblings (that is, adopted siblings) who are reared together are more similar than genetically unrelated siblings who are raised apart.[8]

Biometric Analysis of General Cognitive Ability

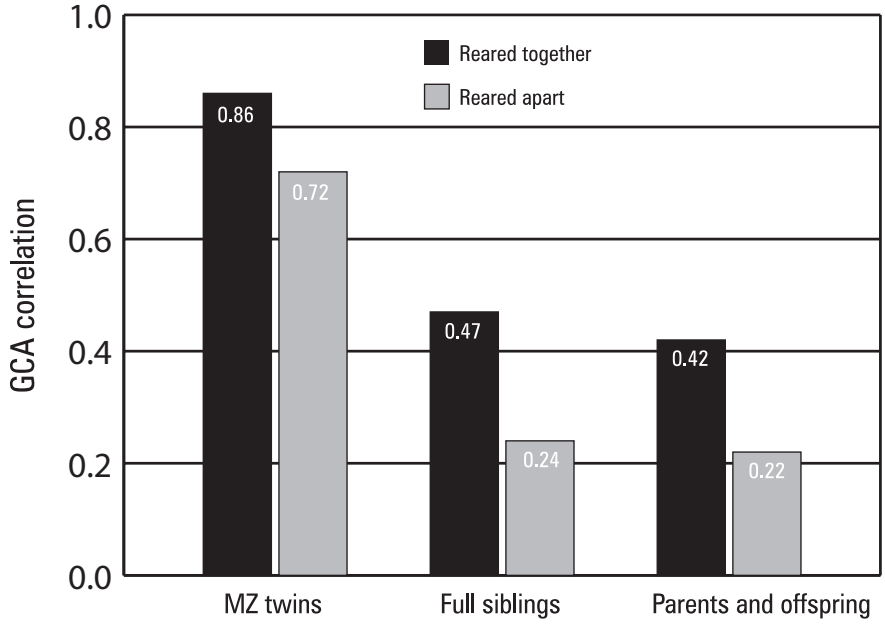
Whereas qualitative comparisons of twin and family correlations suggest that genetic factors and the two types of environmental factors (shared and nonshared) contribute to individual differences in GCA, biometrical genetic methods seek to quantify and formalize these impressions.[9] In the basic biometric model, the phenotypic or observed variance (a measure of the extent to which people differ on the trait in question) is assumed to be an additive combination of three factors: additive genetic effects (A); shared environmental effects (C), which correspond to the influence of factors such as socioeconomic status of the home that reared-together relatives share; and nonshared environmental effects (E), which correspond to the influence of factors such as peer group that differ for individuals reared in the same home. While the basic biometric model is clearly an oversimplification in that it allows for neither gene-environment correlation nor gene-environment interaction, it is generally considered to provide a useful initial

Figure 2.
Average correlation for general cognitive ability among reared-together relatives as a function of degree of genetic relatedness



(T. Bouchard and M. McGue, "Familial Studies of Intelligence: A Review," *Science* 212 [1981]: 1055-59).

Figure 3.
Average correlation for general cognitive ability among reared-together versus reared-apart relatives



(T. Bouchard and M. McGue, "Familial Studies of Intelligence: A Review," *Science* 212 [1981]: 1055-59).

approximation of the relative contributions of three major types of influences on individual differences: A, C, and E. In particular, the proportion of variance associated with genetic factors, which biometricians call the heritability of a trait, is used as an index of how important genetic factors are in explaining why people differ on the trait in question.

There have been three major biometrical analyses of meta-analyzed twin and family correlations for GCA. Table 1 provides a summary of these analyses in terms of the percent of variance in GCA that can be apportioned among the three biometric components, providing an index of the degree to which each component contributes to individual differences in GCA. The results are very consistent across the three sets of analyses: genetic factors are estimated to account for approximately 50 percent of GCA variance, while the nonshared environmental estimate is 14 percent. There is only one small difference among the analyses, and it concerns the shared environmental estimate. In total, C was estimated to account for between 31 percent and 39 percent of GCA variance in the three studies. In a study conducted by Bernie Devlin and colleagues, however, the estimate of C was partitioned into separate prenatal and postnatal contributions, while in the other two analyses only a total estimate was reported.[10] In either case, the biometric models reaffirm the general patterns observed through inspection of the twin and family correlations—both genetics and the shared and nonshared environments contribute to individual differences in GCA.

The consistency of the estimates reported in table 1 might lead some to conclude that the heritability of GCA is 50 percent. Behavioral geneticists recognize, however, that heritability is not a fixed biological constant but that it can instead vary across developmental age, time, and culture. An excellent illustration of the contingent nature of heritability derives from a multinational twin study of reading achievement.[11] In this study, by Stefan Samuelsson and colleagues, the heritability of reading achievement among kindergartners was estimated to be .84, .68, and .33 in Australia, the United States, and Sweden, respectively. The corresponding estimates for the shared environmental effect were .09, .25, and .52. The lower estimate of heritability and higher estimate of shared environment in Sweden versus the other countries was attributed to differences in educational practices. Sweden does not have a formal reading curriculum for kindergartners, while the other countries do. As a consequence, reading progress among Swedish kindergartners depends heavily on exposure at home, a shared environmental effect, while reading progress among kindergartners in the other countries depends more on inherited factors because all students are exposed to a much more homogeneous reading environment in their schools. Confirming this interpretation, when these researchers repeated the study with first graders who received reading instruction at school in all three countries, the heritability of reading achievement varied in a narrow range between .79 and .83, while estimates of the shared environmental effect were all 7 percent or less.

Two factors have been consistently identified as moderating the heritability of GCA. First, as people age, the heritability of GCA increases, while the importance of shared

Variants in the genes underlying GCA may play a role in the mild and idiopathic forms of intellectual disability affecting many individuals. Identifying these genes and learning about that role could eventually contribute to helping such individuals.

environmental influences declines. An analysis of GCA in 10,000 pairs of twins pooled from six separate studies by Claire Haworth and colleagues provides a representative illustration.[12] In this study, twins were classified as being in childhood (ages four to ten), adolescence (ages eleven to thirteen) or early adulthood (ages fourteen to thirty-four), where the broad range of the latter group reflected the design of the individual studies pooled in the analysis. The estimate of GCA heritability increased from 41 percent in the youngest age group to 66 percent in the oldest, while the estimated shared environmental influence decreased from 33 percent to 18 percent. The importance of age moderation is further confirmed in studies of adopted siblings, with an average GCA correlation of .26 in childhood but only .04 in adulthood.[13] The decreasing shared environmental influence on GCA is typically attributed to a diminishing impact of the home environment as individuals age and move away from their rearing homes. Alternatively, the increasing genetic influence has been hypothesized to reflect gene-environment correlational processes; as individuals age and gain increasing control over the nature of their sought-after experiences, they exert that control in a way that reinforces and thus amplifies genetically influenced dispositions.[14]

The second factor found consistently to moderate GCA heritability is the socioeconomic status (SES) of the rearing home. In a landmark study, Eric Turkheimer (a contributing author in this special report) and colleagues explored SES moderation in a sample of young twins from the National Collaborative Perinatal Project.[15] They found that genetic variation explained almost none of the observed variation in the poorest families but helped to explain more than 80 percent of the variation in the richest families. Conversely, shared environmental variation did little to explain observed variation in the richest families but a lot to explain variation (accounting for 60 percent of variation) in the poorest. This pattern of SES moderation has been observed in other, albeit not all, twin studies of GCA.[16] These studies suggest that one effect of an impoverished environment is that it prevents children from achieving their full genetic potential. Further support for this explanation comes from a study by Jeanette Taylor and colleagues, who reported that the heritability of reading achievement was greater among twins taught by highly effective teachers than twins taught by relatively ineffective teachers.[17] Further exploration in large representative samples will be required to inform us about which aspects of SES may be operative and whether the effects observed in children persist into adulthood.

The Search for the Specific Genetic Variants Underlying the Heritability of GCA

As the Human Genome Project progressed in the 1990s, researchers sought to identify the specific genetic variants implied to exist by the heritability studies. Initially, these attempts used "candidate genes": researchers would identify one or more genes that they had reason to believe were relevant to the trait being studied, and they would investigate whether variation in those genes was significantly correlated with the trait of interest. As genotyping became less costly and more efficient, candidate-gene studies proliferated in behavioral and psychiatric genetics. In 2009, Anthony Payton reviewed more than seventy different candidate-gene studies of cognitive function, concluding that the existing literature did not support many earlier reports of significant findings.[18] Christopher Chabris and colleagues investigated twelve genes that, based on their own as well as Payton's review of the literature, they believed had the strongest support for being associated with GCA.[19] In their combined sample of nearly 10,000 individuals, however, none of the variants in these genes was significantly associated with GCA, documenting the general failure of the candidate-gene approach with GCA.

This failure has predictably led to a questioning of heritability studies by critics of behavioral genetics research. Nonetheless, the poor record of candidate-gene studies is not unique to behavioral phenotypes; it is, rather, a general feature of virtually all traits that geneticists have studied.[20] It has been estimated that 95 percent of published findings of associations in human genetics are false positive results, presumably due to some combination of underpowered studies and failure to report and correct for the testing of multiple alternative association models.[21]

Fortunately, an alternative genetic association methodology that takes advantage of remarkable increases in genotyping efficiency has emerged. A genome-wide association study (GWAS) involves genotyping several hundred thousand to several million single nucleotide polymorphisms (or SNPs, changes at a single point in the DNA) located across the entire genome. Unlike a candidate-gene study

that involves a specific polymorphism in a gene that someone has hypothesized is involved in the trait of interest, a GWAS allows for a hypothesis-free search of the entire genome for evidence of genetic association to one or more SNPs. The GWAS approach has resulted in the identification of several thousand genetic associations with several hundred (generally nonbehavioral) clinical phenotypes.[22] Features of the findings from these early GWASs have implications for the feasibility of using this methodology with GCA. Most of the variants identified through GWASs have very small phenotypic effects, with the contribution of any specific locus being much less than 0.5 percent of the genetic variance.[23] To detect true effects that are this small, researchers need to use massive samples. A successful GWAS typically involve consortia with pooled samples of 100,000 or more participants, a formidable challenge for genetic studies of GCA.

Although amassing a sample of more than 100,000 individuals who have been genotyped (on more than one million SNPs) and assessed for GCA may not be feasible in the near term, an alternative, proxy-phenotype approach has been proposed. Rather than study GCA directly, the Social Science Genetics Association Consortium[24] elected to study a credible proxy variable that in modern industrialized societies is highly correlated with GCA: educational attainment (that is, years of completed education). Many behavioral and medical research projects assess educational attainment even if it is not the primary focus of the research, so that the SSGAC was able to create a discovery sample of more than 100,000 individuals and a replication sample of more than 25,000. In their initial GWAS of educational attainment, several significant associations were found,[25] and a subsequent SSGAC study established the validity of the proxy-phenotype approach by showing that three SNPs identified in a GWAS of educational attainment were significantly associated with GCA in an independent sample.[26] As expected, each of these variants accounted for only a very small percentage of GCA variance (about .02 percent). Even though these initial results are modest (the three SNPS combined account for approximately .06 percent of variance in a trait with heritability estimates of 50 percent or more), as the SSGAC continues to build its pooled sample for educational attainment, it is reasonable to expect that the proxy-phenotype approach will produce additional genetic variants associated with GCA. The uphill struggle to link such SNPs to the biology of the brain will constitute the challenge for the current and next generation of behavioral geneticists.

Reactions to the GWAS of educational attainment have been mixed. A lead commentary in *Science* cautiously commended it as an important if somewhat uncertain advance in our understanding of the genetics of general cognitive ability.[27] Other commentators were not as measured or as constructive. The study was dismissed as "unproductive and misleading" in one commentary,[28] while investigations of the genetics of intelligence were said to be "dangerously immoral" in another.[29] Perhaps no area within psychology has received as much ethical scrutiny as genetics research on intelligence,[30] and we agree that such research bears the burden of its early association with the eugenics movement.[31] Nonetheless, as the Nuffield Council on Bioethics concluded in its comprehensive review of behavioral genetics, research like that in which contemporary behavioral geneticists engage helped to undermine the eugenics movement by proving that eugenic policies could not achieve their stated goals.[32] When ethical analyses have considered not only the potential for research on the genetics of intelligence to do harm but also its potential to provide benefits, the balance has fallen in favor of continued research in this area.[33]

1. C. Spearman, "General Intelligence: Objectively Determined and Measured," *American Psychologist* 15 (1904): 201-93.

2. A. R. Jensen, *The G Factor: The Science of Mental Ability* (Westport, CT: Praeger, 1998).

3. D. Lubinski, Introduction to the Special Section on Cognitive Abilities: 100 years after Spearman's (1904) "'General intelligence,' Objectively Determined and Measured," *Journal of Personality and Social Psychology* 86, no. 1 (2004): 96-111.

4. L. S. Gottfredson, "Why *g* Matters: The Complexity of Everyday Life," *Intelligence* 24, no. 1 (1997): 79-132; L. S. Gottfredson and I. J. Deary, "Intelligence Predicts Health and Longevity, But Why?," *Current Directions in Psychological Science* 13 (2004): 1-5.

5. J. A. F. Roberts, "The Genetics of Mental Deficiency," *Eugenics Review* 44, no. 2 (1952): 71-83; E. Zigler and R. M. Hodapp, *Understanding Mental Retardation* (New York: Cambridge University Press, 1986).

6. T. J. Bouchard and M. McGue, "Familial Studies of Intelligence: A Review," *Science* 212 (1981): 1055-59.

7. R. Plomin and I. J. Deary, "Genetics and Intelligence Differences: Five Special Findings," *Molecular Psychiatry* 20, no. 1 (2015): 98-108.

8. T. J. Bouchard, "Genetic Influence on Human Intelligence (Spearman's *g*): How Much?," *Annals of Human Biology* 36, no. 5 (2009): 527-44.

9. M. C. Neale and L. R. Cardon, *Methodology for Genetic Studies of Twins and Families* (Dordrecht, Netherlands: Kluwer Academic, 1992).

10. B. Devlin, M. Daniels, and K. Roeder, "The Heritability of IQ," *Nature* 388 (1997): 468-71.

11. S. Samuelsson et al., "Response to Early Literacy Instruction in the United States, Australia, and Scandinavia: A Behavioral-Genetic Analysis," *Learning and Individual Differences* 18, no 3 (2008) 289-95.

12. C. M. A. Haworth et al., "The Heritability of General Cognitive Ability Increases Linearly from Childhood to Young Adulthood," *Molecular Psychiatry* 15, no. 11 (2010): 1112-20.

13. Bouchard, "Genetic Influence on Human Intelligence (Spearman's *g*)."

14. S. Scarr and K. McCartney, "How People Make Their Own Environments: A Theory of Genotype-Environment Effects," *Child Development* 54, no. 2 (1983): 424-35; D. A. Briley and E. M. Tucker-Drob, "Explaining the Increasing Heritability of Cognitive

Ability across Development: A Meta-Analysis of Longitudinal Twin and Adoption Studies," *Psychological Science* 24, no. 9 (2013): 1704-13.

15. E. Turkheimer et al., "Socioeconomic Status Modifies Heritability of IQ in Young Children," *Psychological Science* 14, no. 6 (2003): 623-28.

16. T. C. Bates, G. J. Lewis, and A. Weiss, "Childhood Socioeconomic Status Amplifies Genetic Effects on Adult Intelligence," *Psychological Science* 24, no. 10 (2013): 2111-16.

17. J. Taylor et al., "Teacher Quality Moderates the Genetic Effects on Early Reading," *Science* 328 (2010): 512-14.

18. A. Payton, "The Impact of Genetic Research on our Understanding of Normal Cognitive Ageing: 1995 to 2009," *Neuropsychology Review* 19, no. 4 (2009): 451-77.

19. C. F. Chabris et al., "Most Reported Genetic Associations with General Intelligence Are Probably False Positives," *Psychological Science* 23, no. 11 (2012): 1314-23.

20. J. P. A. Ioannidis et al., "Replication Validity of Genetic Association Studies," *Nature Genetics* 29, no. 3 (2001): 306-9.

21. H. M. Colhoun, P. M. McKeigue, and G. D. Smith, "Problems of Reporting Genetic Associations with Complex Outcomes," *The Lancet* 361 (2003): 865-72.

22. P. M. Visscher et al., "Five Years of GWAS Discovery," *American Journal of Human Genetics* 90, no. 1 (2012): 7-24.

23. J. H. Park et al., "Estimation of Effect Size Distribution from Genome-Wide Association Studies and Implications for Future Discoveries," *Nature Genetics* 42, no. 7 (2010): 570-75.

24. See Social Science Genetic Association Consortium, http://www.ssgac.org/.

25. C. A. Rietveld et al., "GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment," *Science* 340 (2013): 1467-71.

26. C. A. Rietveld et al., "Common Genetic Variants Associated with Cognitive Performance Identified Using the Proxy-Phenotype Method," *Proceedings of the National Academy of Sciences of the United States of America* 111, no. 38 (2014): 13790-94.

27. J. Flint and M. Munafo, "Herit-Ability," *Science* 340 (2013): 1416-17.

28. D. Graur, Genes and Education: A Summary of a #GWAS Article in Twelve Languages, 2013, http://judgestarling.tumblr.com/post/54236053456/genes-and-education-a-summary-of-a-gwas-article.

29. A. Buchanan, "The Significance of (Looking for) Genes for Educational Achievement," *The Mermaid's Tale* (blog), July 8, 2013, http://ecodevoevo.blogspot.com/2013/07/the-significance-of-looking-for-genes.html.

30. E. Parens, A. R. Chapman, and N. Press, *Wrestling with Behavioral Genetics: Science, Ethics, and Public Conversation* (Baltimore, MD: Johns Hopkins University Press, 2006).

31. I. I. Gottesman and L. Erlenmeyer-Kimling, "Prologue: A Foundation for Informed Eugenics," *Social Biology* 18, supplement (1971): S1-S8; M. McGue, "The End of Behavioral Genetics?" *Behavior Genetics* 40, no. 3 (2010): 284-96.

32. Nuffield Council on Bioethics, *Genetics and Human Behaviour: The Ethical Context* (London: Nuffield Council on Bioethics, 2002).

33. A. Newson and R. Williamson, "Should We Undertake Genetic Research on Intelligence?," *Bioethics* 13, nos. 3-4 (1999): 327-42; M. J. Reiss, "The Ethics of Genetic Research on Intelligence," *Bioethics* 14, no. 1 (2000): 1-15.