Chapter 9
Behavior-Genetic Studies of Academic Performance in School Students: A Commentary for Professionals in Psychology and Education

Brian Byrne, Richard K. Olson and Stefan Samuelsson

Abstract Available behavior-genetic research indicates that the single largest factor influencing individual differences in literacy development is genetic endowment. We briefly review some typical evidence and methodology used in studying the behavior-genetics of reading. We then outline three hypothetical educational scenarios and demonstrate how behavior-genetic studies might play out in them, with the aim of enhancing the critical capacity of school psychologists and other educational professionals to evaluate research findings in this area. We show that heritability estimates will tend to be higher in educational environments in which the instruction and other factors are more uniform, that the way subsamples are combined can affect estimates, and that population-level estimates cannot be used to determine the etiology of any individual child's performance. We address and dismiss genetic determinism, and review evidence to suggest that genetic accounts of reading disability may reduce blame and stigma yet increase pessimism about successful intervention. However, we argue that continued research into optimal ways to design and deliver curricula is quite compatible with the substantial heritability of individual differences in literacy and has already provided grounds for optimism. We also suggest that genetically derived constraints on academic progress bring into sharp focus questions about the goals of education.

Differences among students in reading and spelling are matters of great interest to educators and are the subjects of substantial amounts of research. Prominent among the factors that affect these differential aspects of achievement is the student's genetic endowment. The goal of this essay is to help psychologists and educators to enhance their understanding of what this fact means, and does not mean, for policy and

B. Byrne (✉)
University of New England, Armidale, Australia
e-mail: bbyrne@une.edu.au

R. K. Olson
University of Colorado, Boulder, USA

S. Samuelsson
Linköping University, Linköping, Sweden

D. A. Kilpatrick et al. (eds.), *Reading Development and Difficulties*,
https://doi.org/10.1007/978-3-030-26550-2_9

practice. We try to achieve this goal in part by considering how genetically sensitive research methods may play out in a range of educational scenarios. We link these hypothetical situations to real data where these exist.

It is not our aim to provide a systematic review of findings about the relative influence of genes and the environment on literacy. Such reviews are available (e.g., Asbury & Plomin, 2013; Kovas, Haworth, Dale, & Plomin, 2007; Little, Haughbrook, & Hart, 2017; Olson, Keenan, Byrne, & Samuelsson, 2014). Nevertheless, we offer some examples of these kinds of studies to (a) justify the claim that genes play a substantial role in literacy development, (b) lay a foundation for further appreciation of the methods employed in this area of research, and (c) consider the consequences for policy and practice of acknowledging the reality of the substantial genetic influence on literacy development.

9.1 Methods

Most *behavior-genetic* research, as it is typically termed, uses twins. Research of this sort takes advantage of the fact that there exist monozygotic (*MZ*, or "identical") and dizygotic (*DZ*, or "fraternal") twin types. Members of a monozygotic pair share all of their genes and members of a dizygotic pair share, on average, half of their segregating genes (i.e., genes that make individuals different from each other). When twins are raised together, as is overwhelmingly the case, *heritability* can be estimated as twice the difference between the within-pair monozygotic and dizygotic correlations. For example, if the correlation between monozygotic twins' performance on a reading test is X, and the correlation between dizygotic twins' performance on that same test is Y, then heritability is estimated at the difference between X and Y times two. *Variance* is a statistical way to measure the spread of the variation or variability (e.g., high-reading skills to low-reading skills) within a group. Heritability refers to the proportion of the variance on the trait that can be attributed to additive genetic variance. The proportion of the trait's variance that is not heritable is attributable to environmental factors that affect values on the trait (e.g., scores on reading tests). Twin studies allow for those factors to be partitioned into (1) ones that similarly influence both twins in a pair (i.e., *shared environment* such as family socioeconomic status, schools attended, and common teachers when both twins share schools and teachers), and (2) ones that twins do not share (i.e., *non-shared environment* such as individual illnesses, accidents, and separate teachers and friends). These values are generally calculated using a version of a statistical methodology called structural equation modeling that has been specially designed for twin data. This allows for computation of confidence intervals and tests of the statistical significance of values (Neale, Bokor, Xie, & Maes, 2002).

The validity of conclusions from classic twin design relies on several assumptions. One assumption is the *equal environments assumption* under which it is assumed that monozygotic twins are not treated more similarly than dizygotic twins. A second assumption is that there is no *assortative mating* for the trait under investigation;

for example, that adults do not select their mate based on similar levels of a trait. Violations of these assumptions can affect the accuracy of estimates of heritability and environmental influence unless the violations can be accommodated within the statistical modeling. There is a substantial literature on these and other assumptions, and we recommend a recent review by Barnes et al. (2014), which shows that twin research is generally robust in the face of possible violations of these assumptions (see also Plomin, DeFries, Knopik, & Neiderhiser, 2013).

9.2 Some Results

In Table 9.1, we document results from a range of studies using twin methodology. Just the heritability (A) and shared environment components (C) of variance are included in the table. The unique environment (E) is the remainder in each case (i.e., elements of the environment are not shared within a twin pair). Thus, E = 1 − (A + C). The values are for individual differences except when identified as member of the lower end of the skill distribution (lowest 10%). When *latent variables,* such as reading comprehension, are indicated in the table, such variables are created by using multiple similar measures (e.g., different reading comprehension tests) and modeled as latent variables (i.e., the performance overlap across multiple reading comprehension tests provides a better estimate of reading comprehension than a single test, reducing measurement error).

It should be clear from Table 9.1 that the consistent finding is that genetic endowment accounts for a substantial proportion of the variability we see among children in their reading and spelling skills. This also includes whether or not a student performs at the bottom end of the distribution for reading and spelling, often referred to as *dyslexia* (Snowling & Hulme, 2005). The minimum heritability value is .39 and the maximum is .87. The studies vary somewhat in the contribution of shared environment to variance, and possible reasons for this are discussed later in the chapter.

We want to focus in particular on data from the *International Longitudinal Twin Study*, which is a longitudinal study of literacy development across four countries and three languages. We have selected this study not primarily because it is one that we (the authors) initiated but because its international component allows us to illustrate some interesting features of twin research.

Twin children in Australia, Norway, Sweden, and the USA were enrolled and assessed in their final preschool year (aged 4–5) and followed in literacy development for at least three further years near the end of kindergarten first grade and second grade. The total number of twin pairs was approximately 1000, with around half from the USA and a quarter each from Australia and Scandinavia (Norway and Sweden). They were assessed on a wide variety of literacy, cognitive, behavioral, and familial characteristics, but the focus here is on literacy and its known precursors. Some of these findings were presented in Table 9.1. Across several school grades, the results show: (1) a high level of heritability for multiple early literacy skills, (2) minimal shared environmental effects, and 3) modest unique environmental effects. It must

Table 9.1 Example results from twin studies of literacy

Sample	Measure	A	C	Reference
ILTS Kindergarten	Word identification	.70	.22	Byrne et al., (2010)
ILTS Kindergarten	Spelling	.39	.40	Byrne et al., (2010)
ILTS Grade 1	Word identification	.83	.01	Byrne et al., (2010)
ILTS Grade 1	Reading comprehension	.76	.03	Byrne et al., (2010)
ILTS Grade 1	Spelling	.72	.06	Byrne et al., (2010)
ILTS Grade 4	Word identification	.77	.14	Olson et al., (2011)
ILTS Grade 4	Reading comprehension	.86	.09	Olson et al., (2011)
NAPLAN Grade 3	Reading comprehension	.71	.05	Grasby et al., (2016)
NAPLAN Grade 9	Reading comprehension	.61	.13	Grasby et al., (2016)
CLDRC Multiple grades	Word identification LV	.87	.10	Christopher et al., (2016)
CLDRC Multiple grades	Reading comprehension LV	.82	.18	Christopher et al., (2016)
CLDRC Multiple grades, group membership, lowest 10%	Reading and spelling	.61	.30	Olson et al., (2014)
Florida Grade 1	Reading fluency	.62	.22	Taylor & Schatschneider, (2010)
TEDS Male 7-year-olds, group membership, lowest 10%,	Word identification	.67	.21	Harlaar et al., (2005)
TEDS Female 7-year-olds, group membership, lowest 10%,	Word identification	.50	.40	Harlaar et al., (2005)
WRRP Kindergarten + Grade 1	Word identification	.55	.34	Petrill et al., (2007)

Note: *A* Additive genetic influence; *C* Shared environment influence; *ILTS* International Longitudinal Twin Study; *CLDRC* Colorado Learning Disabilities Research Center; *NAPLAN* National Assessment Program: Literacy and Numeracy (Australia); *WRRP* Western Reserve Reading Project; *TEDS* Twins Early Development Study; *LV* Latent Variable

be noted that any measurement error is included in the figure representing the unique environmental effects.

9.3 The Story Becomes More Complex

As a way of introducing the contrasting scenarios we wish to employ, consider results from Samuelsson et al. (2008). They reported on reading test scores at the end of the first school year, kindergarten, separately for each country (combining Norway and Sweden in a single "country," Scandinavia). The scores were based on a combination of word and nonword reading efficiency from the *Test of Word Reading Efficiency* (Torgesen, Wagner & Rashotte, 1999). In Australia, the estimates of genetic, shared environment, and nonshared environmental influences were .84, .09, and .08, respectively. In the USA, they were .68, .25, and .07, and in Scandinavia, they were .33, .52, and .15. The contrast between Australia and Scandinavia is particularly marked, with the Australian heritability estimates being over two-and-a-half times higher and the estimates of shared environment being over five times lower than in Scandinavia. The U.S. estimates sit between these extremes.

It is unlikely that genetic differences between the nations account for these differing patterns of heritability because at the end of the second school year, the three countries fall into line, with almost identical heritability estimates of .80, .83, and .79 for Australia, the USA, and Scandinavia, respectively. Instead, it is likely that differing educational environments explain the contrasting results. In New South Wales, Australia (NSW., the site of the sample), kindergarten children attend a full school week (9 a.m. to 3 p.m.) and are subject to a state-wide curriculum that mandates that 35% of the time is spent on literacy instruction with agreed benchmarks as achievement targets throughout the year. In Colorado, the site of the U.S. sample, attendance is limited to half days (3–4 hours), and there is no state-mandated curriculum for kindergarten literacy, likely implying less uniformity in quantity and quality of instruction. In Scandinavia (at least at the time of the research), kindergarten attendance is not compulsory, although almost all children do attend, and the emphasis is on social and emotional development, with any literacy instruction being given informally and mostly at home. Thus, it appears that across the three sites, there is a continuum of intensity and extent of literacy instruction that corresponds to a continuum of heritability, which in turn is traded off with shared environmental influence.

A second result with similar implications arose during the course of the International Longitudinal Twin Study when about halfway through the multiyear recruitment process, Norway introduced formal literacy instruction into kindergarten, with a dedicated 6.2 hours per week. The educational change occurred in 2007. If the idea that the differences in heritability between Scandinavia and Australia just outlined were due to instructional differences, one would expect to see a change in the heritability estimates (and in average literacy scores) in the Norwegian sample exposed to the new, literacy-focused curriculum in comparison with those twins exposed to the previous curriculum. This is indeed what appears to have happened (Samuelsson, Byrne, Hulslander, & Olson, 2009). Word and nonword reading scores more than doubled, and spelling scores increased by almost 50%. The heritability of reading changed from .32 to .40, with a corresponding drop in the shared environmental

effect. The spelling results were clearer with heritability increasing from .44 to .72, and shared environment dropping to .05. The twin numbers were (unavoidably) too small at 102 and 61 pairs, respectively for the old and new curricula, for significance testing, but the direction of the changes is exactly what would be expected on the hypothesis that increasing the intensity and uniformity of literacy instruction will lead to increased heritability estimates. These results are, by the way, particularly compelling evidence against the idea that genetic differences among the countries studied are behind the changes in heritability that are generated by changes in curricula.

The broad lesson from this set of results is that it is not appropriate to speak of *the* heritability of some variable. It is better to speak of the heritability of a variable *under X circumstances* (of sample, environmental particularities, period in history, and so on). And continuing the lesson is the greater the environmental range (as in Scandinavia for kindergarten literacy instruction), and the greater the size of the environmental influence. Heritability will tend to be highest in relatively uniform educational environments in which all children receive similar literacy opportunities (see also Asbury and Plomin, 2013, on this important point).

Armed with these concepts, let us now consider how behavior-genetic research might play out in some hypothetical educational systems with a view to refining our understanding of the implications of such research.

9.4 Scenarios

9.4.1 Scenario 1

The situation we consider here is: *Universal free education with mandatory attendance, a centralized curriculum, and a teaching workforce that delivers the curriculum in a uniform way within uniform school structures*.

This scenario incorporates minimal school environmental variability. In comparison with the situation for literacy instruction in kindergarten in Scandinavia and the USA., the NSW, educational jurisdiction in Australia appears to come closer to this scenario. Here, heritability was higher than in the other sites in the International Longitudinal Twin Study, and at .84, quite high indeed. Environmental influences that stem from school factors, such as the influence of individual teachers and class size, will decline to a negligible level.

Note that there is no information about how *well* students are performing. The curriculum may be suboptimal or teacher preparation might be flawed with all teachers trained to teach to the same (low) standard. That is, high heritability under these circumstances could go hand in hand with a low-jurisdiction-wide performance in comparison with other jurisdictions.

Individual children in Scenario 1: Even if heritability is high in the sampled population, this cannot be taken to mean that in any individual case of low-academic

achievement, genes are the cause. The roughly 20–50% of population variance that is not attributable to genes leaves ample room for environmental factors to affect achievement in individual cases. Further, in the current state of (lack of) knowledge about actual genetic markers for low-level performance (discussed later), we have no way of identifying which children's problems are primarily genetic in origin and which are not. But even in the cases of (those unknown) children whose difficulties are primarily genetic in origin, well-designed and well-delivered intervention is likely to foster their literacy development (again, more later). Thus, the search for how individual environments affect literacy and its development remains a fully justified research endeavor even in the face of substantial group heritability in particular educational jurisdictions.

Groups of children under Scenario 1: Importantly, there may still be *groups* within the jurisdiction whose poor group performance is not attributable to genes. In Australia, for example, school attendance by indigenous students lags up to 20% behind nonindigenous students and (probably not coincidentally) their literacy performance lags as well (Purdie & Buckley, 2010). However indigenous Australians comprise only about 2.5% of the population, and so even if a behavior-genetic study was thoroughly representative of the population, which is in any case unlikely, the environmental circumstances of this small but important group is likely to have an undetectable influence on the estimate of the shared (or nonshared) environmental effect derived from twin studies.

As a second example, it has been reported that literacy (and numeracy) achievement across several grades assessed in nation-wide tests in Australia is adversely affected by arsenic, cadmium, and lead contamination in soil, dust, and aerosols in three cities with long histories of mining and metals processing (Dong, Taylor, Kristensen, & Zahran, 2015). The effect was magnified after controlling for school SES levels.

The general point emerging from these two examples is that high heritability should not inhibit the search for environmental causes of variance that apply to groups of students. Population subgroups, particularly relatively small ones, subject to deleterious environmental influences can remain below the radar in behavior-genetic studies using national samples.

Trends toward educational uniformity: Insofar as a nation, a state, or a school district may move toward a common curriculum for some or all school subjects, it can be expected that genes will become more prominent as a determinant of individual differences in academic achievement at the national (or state or school district) level. The same holds for moves toward uniform training standards for teachers and for aspects of school organization, such as class sizes. As a consequence, educational authorities who support these trends will need to accept an increasing role for genes in the variability of student achievement.

9.4.2 Scenario 2

As opposed to Scenario 1, here, we consider: *Elective education, no state financial support, with minimal guidelines for curriculum and its delivery in schools that vary in organizational structure*.

The picture here is education only for families that can afford to send their children to school, with few constraints on design and delivery of instruction. By the way of example, some of these characteristics are largely true of education in the African country of Swaziland. In Swaziland, the costs of education are prohibitive for many families. According to a World Bank report (Marope, 2010), 16% of Swaziland children were not enrolled in primary school, 74% were not enrolled in junior high school, and 88% were not enrolled in senior high school. The report also indicates that "school curricula do not clearly stipulate the skills and competencies that learners should acquire at each level" (p. 68) and states that teacher quality is subject to only scant checks. We are not saying that Swaziland is unique in its educational profile or that it should be criticized for it—after all, the country is burdened with very high financial and social costs of HIV/AIDS and has few natural resources other than agriculture and forestry. We simply wish to show that Scenario 2 is realistic.

A twin study of school achievement in the kind of environment is that Scenario 2 entails would, we hypothesize, show minimal genetic influence and substantial environmental effects, assuming that little in the way of informal education is offered in the home. These effects would be large of the shared environment kind if both twins in a family either attended school or did not attend school, and with increasing non-shared environmental effects if families could only afford to send one twin of a pair to school. The environmental effects would stem from the fact that zygosity would matter less for the degree to which twins are alike (the engine of etiological estimates in twin research) compared to the contrast between schooled and unschooled children. In other words, the contrast between monozygotic and dizygotic twin similarity would be dwarfed by the contrast between schooled and unschooled twin similarity.

This example illustrates that it would be a mistake to extrapolate to the world stage that the consistent findings of substantial heritability of literacy development derived from research in western societies. Access to education remains the major concern in many nations, and researchers who are working to bring to these countries affordable literacy lessons using scientifically validated methods deserve continuing support. For an account of one such project with a focus on African nations, initiated by Lyttinen and colleagues in Finland, see Ojanen et al. (2015). It is encouraging that UNESCO is supporting efforts like this (see https://www.jyu.fi/en/news/archive/2015/01/tiedote-2015-01-29-11-23-14-298438).

9.4.3 Scenario 3

Here, we consider: *Hybrid systems*.

Educational systems are in reality complex structures and may be getting more complex with a push, seen in some nations, to devolve more responsibility to individual schools for matters such as finance, staffing, curriculum and textbook choice, and management of special-needs students. In North America, giving states and provinces more control over educational decision making has a history going back to the 1970s in Florida and Alberta (Australian Education Union, 2012), and for a considerable number of years, legal and other issues surrounding this practice have been on the agenda (e.g., Florestal & Cooper, 1997; though a push for greater national common-core standards has characterized the last several years in the USA). So imagine, hypothetically, that within an educational jurisdiction that maintained overall responsibility for education, say NSW in Australia, there was considerable variety in curricula across individual subdistricts. Imagine further that some curricula were of high quality and some were quite suboptimal and that this difference influenced student performance. What would be the consequences for estimates of genetic and environmental influences?

The answer depends on the question being asked and on how it is answered. If the question is, "How heritable is literacy in Grade 2 in (say) N.S.W.?" then district-by-district variability in performance as a function of curriculum quality would increase the shared environmental effect, trading off with heritability. This is because, as a genuine factor affecting performance in children in general, the effect of the curriculum would be to raise or lower twins' scores irrespective of zygosity, thereby reducing the contrast between monozygotic and dizygotic twin correlations compared to a situation when there were uniformly effective curricula, or just one. In contrast, if the question is, "How heritable is literacy in District A, B, and so on," curriculum disappears as a factor, meaning that the shared environment influence will be limited to other things, such as the home and peers.

To the issue of how the question is answered and returning to our first one–heritability in NSW—it is common practice in behavior-genetic research to standardize within variables of no or minimal interest to the question in focus. For example, differences in age within a group of second-grade children may affect the scores on a literacy test. But if interest is in heritability in second grade irrespective of age, for example, when comparing with heritability in fourth grade, age will be controlled for statistically. The same would be true for, say, gender if it were not the focus of a particular study. Now, if in the hypothetical study of heritability in NSW researchers standardized literacy scores *within sub-districts*, the curriculum effect would disappear as an influence because all subdistricts would have means of zero (and *SD*s of unity) and heritability would be estimated as higher than when they did not do so. It is easy to scale-up this issue, say to the entire USA. If there was state-based effective environmental variability, for example, in teacher qualification standards (Darling-Hammond, 2000), standardizing within states would obscure this fact and

simultaneously increase heritability estimates above those that would be obtained had standardization not been implemented.

Some data are available on this contrast (Byrne, Olson, & Samuelsson, 2013). As mentioned, the heritability of reading at the end of Grade 1 was almost identical in Scandinavia and Australia, at .80 and .79, respectively (with the USA at .83). But the Australian twins were reading at higher levels than their Scandinavian counterparts, with respective means of 43.8 ($SD = 16.3$, $N = 502$) and 25.6 (14.6, 576) on the identical (translated) test of word reading that was used. Similarly, mean reading comprehension scores were very different in the two samples, 26.1 ($SD = 7.4$) and 13.5 (8.3). It can be presumed that the inequalities come about because the Australian twins had just completed two years of formal reading instruction in comparison with the one year in Scandinavia, a clear environmentally driven difference. For word reading, the genetic and shared environmental influences when the two datasets are combined *after* within-country standardization are .83 and .00. In contrast, when the datasets are combined *without* within-country standardization, heritability drops to .59 and shared environment rises to .29. The same shift happens when reading comprehension is the literacy variable employed: following within-country standardization, .62 and .16; lacking within-country standardization, .38 and .49.

Earlier, we cautioned against referring to *the* heritability of a variable, instead referring to its heritability under X circumstances. To this, we can add, and *using Y analytic choices*.

9.5 Interim Summary and Looking Ahead

We have presented a variety of issues that educational professionals should be aware of in interpreting research that quantifies the relative influence of genes and various aspects of the environment on literacy development. They mostly boil down to being cognizant of the research settings and methodologies and realizing that if these change, so may the estimates. In particular, the trade-off between heritability and environmentality according to the environmental circumstances needs to be acknowledged. The more restricted and uniform the environmental range, the higher the heritability is likely to be. The more varied this environmental range, the greater is the potential for environmental influence. In some circumstances, subgroups in a population subject to the special effects of the environment may be invisible in national samples (e.g., an indigenous population in Australia, mentioned above). Finally, computational choices may affect the quantitative estimates, with our example being the choice to standardize or not standardize subsamples within a country or region. Such standardization within subsamples may obscure environmental effects that are real.

Despite these caveats, the weight of evidence is for a substantial effect of genes on literacy development, at least within those individual educational environments that have been studied so far. In the remainder of the chapter, we will discuss some of the implications of this conclusion. We first discuss the pitfalls of "genetic determinism,"

and then illustrate (a) how evidence for substantial heritability can be something of a mixed blessing, (b) how further research within the behavioral genetic framework may be able to guide instruction and intervention for children who struggle with reading, and (c) how the findings fit into the broader issue of the goals of education.

9.6 Genes Matter, But …

For professionals dealing with children and (and adults) struggling with literacy, and for the families of these children (and adult sufferers themselves), an understanding of how genes influence behavior is important. Here are some observations:

- There is no single gene responsible for reading (dis)ability. Despite media headlines that sometimes declare that *the* gene for X has been discovered, monogenetic disorders are extremely rare (Chabris, Lee, Cesarini, Benjamin, & Laibson, 2015). Behavioral traits are usually affected by thousands of individual mutations, each of tiny effect; for example, it has been estimated that schizophrenia, which is substantially heritable, is associated with 8300 "SNPs," *single-nucleotide polymorphisms*. It is likely that academic achievement skills are affected by even more (Chabris et al., 2015).
- The environment has, or can have, a substantial role in whether and how genes influence traits. A good example of the "can have" part is successful dietary manipulation and supplementation in cases of the metabolic disorder phenylketonuria. In the cases of marked reading disability, there is evidence that preventive efforts and well-designed and timely intervention can have long-lasting positive effects on literacy development in children showing signs of risk for reading disability (e.g., Blachman, Schatschneider, Fletcher, Murray, & Munger, 2014; Elbro & Petersen, 2004; Kilpatrick & O'Brien, Chap. 8, this volume).
- The environment can amplify the influence of genes through gene environment correlation (Plomin et al., 2013). In the case of literacy, children whose genes support reading skills tend to read more, thereby accruing the benefits of further reading practice. In contrast, children whose genes lead them to struggle with reading will in all likelihood read less, missing out on the benefits of extensive reading experience (for a summary of evidence supporting these expectations, see Olson et al., 2014). But although this may appear to be unalloyed bad news for those less well endowed genetically for reading, it also means that if these children *can* be encouraged to read, then what is essentially an environmental effect (reading experience) can minimize the amplification effect, particularly when accompanied by effective intervention opportunities (see Chap. 8). We return to the role of reading experience later.
- Epigenetics further complicates any simple picture of genetic influence. These are biochemical processes that leave DNA sequences intact but alter the expression of genes for good or ill (see Carey, 2012 for an accessible introduction). The environment, in the form of such factors as toxins, dietary abnormalities, and stress,

can modify gene expression through methylation and acetylation of nucleotides and histones. Among other things, this can mean that children with the same alleles at the sites of genes that are candidates for reading difficulties may be differentially affected.

Thus, genes cannot be conceptualized as deterministic (for an extended discussion, see Dar-Nimrod and Heine, 2011). Their possible effects are subject to opportunistic and deliberate environmental influences, and there is ample room for complex interactions among the many genes that affect a trait. The environmental influences in the case of literacy include, of course, preventative and remedial interventions. In Chap. 8, studies are reviewed that document improvements and, importantly, the maintenance of these improvements, even among students with the most significant word-level reading problems. Not all interventions produce large improvements or ones that last when followed up across months and/or years, but in Chap. 8, the principles that distinguish more from less impressive treatments are proposed. This should encourage practitioners to implement the quality interventions described in Chap. 8, despite the well-established findings that genes play a substantial role in accounting for variation in reading acquisition

The fact that many genes and their interactions are involved in reading development should discourage practitioners from looking forward to the day when a simple "genetic test" will accurately identify children at risk for reading disability. There is no realistic prospect of that happening.

9.7 Mixed Blessings

It is important for professionals and others to appreciate those "biogenetic" accounts of behavioral disorders are, in the words of Haslam and Kvaale (2015), a mixed blessing. Kvaale, Haslam, and Gottdeiner (2013) conducted a meta-analysis of what they refer to as the "medicalization" of psychological problems. Although reading disability was not among the disorders included in the research they summarized, and although we know of no research that has systematically studied literacy problems or other learning difficulties in this context, it is reasonable and prudent to assume that the conclusions that they draw would apply to these school-based disabilities.

The conclusions of Kvaale et al. (2013) were that providing people with biogenetic explanations (a) reduced stigma and blame associated with the disorders, but (b) increased pessimism about prognosis. In the authors' words, and in the context of mental health, "[m]edicalization of these problems may reduce blame, but at a cost. For example, pessimism about change could take hold of the affected individual, family members, mental health professionals, and the society at large, setting the stage for self-fulfilling prophesies that could seriously impede the recovery process" (p. 790). If these conclusions do indeed transfer to underachievement in reading, we could envisage relief from blame accompanied by a kind of fatalism on the part of parents, classroom teachers, and remediation specialists about prospects for recovery.

It is indeed possible that apprehension about a developing fatalism in part motivates the documented reluctance of many individuals working in education to acknowledge genetic influences on school achievement (Grigorenko, 2007).

Kvaale et al. (2013) underline the importance of not misinforming the public about biogenetic explanations, but urge caution in doing so to avoid negative side effects. The educational community needs to develop a narrative that reflects the scientific findings of reading disability but that at the same time discourages pessimism in affected individuals, their families, and relevant professionals. The International Dyslexia Association (IDA) is one organization that has faced the dilemma squarely. Its website (http://www.interdys.org) declares that dyslexia has neurobiological and genetic roots and at the same time encourages optimism that the difficulties can be overcome with proper diagnosis, appropriate phonologically based and multisensory training, "hard work," and the support of family, friends, and teachers. They identify successful figures in the arts, science, entertainment, and other fields who are classed as "dyslexic."

As helpful as the IDA's stance may be to those who engage with it, we are not aware of similar messages circulating in mainstream educational circles. With the growth of state- and nation-wide assessment in the basic school subjects of literacy and numeracy, and with scores made available in ways that families can compare their children with others and against criteria of achievement for each grade, many families will now discover low levels of performance in their children for the first time. If simultaneously public awareness of biogenetic accounts of school achievement increases, the unfortunate side effects that Kvaale et al. (2013) document may also become pervasive. The alternative of suppressing the biogenetic "story" is clearly unacceptable, and so the development of a compensating narrative becomes urgent in our view. Some of the highly encouraging results regarding the prevention and intervention with reading disabilities described in Chap. 8 can provide a basis for such a narrative.

9.8 Using Behavior-Genetic Studies to Inform Instruction and Intervention

It is one thing to quantify the relative influences of genes and the environment on a trait such as literacy. It is quite another thing to identify *how* these factors separately and together determine levels of the trait. We should note at the start that we are still unable to say much about the molecular biology responsible for variability even though there has been some progress at that level of explanation for severe forms of reading disability, with a handful of genes associated with the disorder being identified in replicated studies (Poelmans, Buitelaar, Pauls, & Franke, 2011). The main problem with gene identification for complex human characteristics is that a vast gap remains between the genetic variability that known genes can account for and the genetic variability itself based on behavioral studies such as those with

twins. In medicine and other domains, this has been called the *missing heritability problem*, where typically liability for a disease that can be derived from identified genetic variants is just a few percent of the actual genetic liability (Maher, 2008). It is in fact this gap that forms the basis, mentioned earlier, for believing that complex traits are the product of very many genes, each of very small average effect in the population (Chabris et al., 2015).

Nevertheless, even though we know little about the actual genes affecting literacy, well-designed research can get closer to underlying cognitive processes than studies that simply analyze literacy itself. Measures of these cognitive processes need to be included alongside measures of literacy, and then, through multivariate genetic analyses, researchers can determine whether the processes underlying these measures are genetically correlated with the literacy phenotype. A common-sense analysis of learning to read words (one backed by a substantial research literature—see chapters in the volume edited by Snowling and Hulme [2005]) would implicate as a minimum (a) learning processes that can bind graphic symbols with morphophonological elements (initially, letters and phonemes, and later, letter strings and whole words), and (b) processes that support left-to-right decoding of letter strings, where a child needs to recover the phoneme that the first letter represents, hold it in mind while recovering the next letter's phoneme, continue this process, and finally, amalgamate the phoneme string into a word. The first of these is generally referred to as associative learning, the second as working memory.

In the International Longitudinal Twin Study, the researchers included several tasks that tap associative learning for graphic symbols and working memory for linguistic material. They indeed discovered that they are genetically correlated with (i.e., are *pleiotropic* with) early letter knowledge and later word reading efficiency (Byrne, Wadsworth, et al., 2013).

These findings immediately suggest diagnostics to predict whether a child will struggle with learning to read and remedial steps for those who subsequently do. Early difficulties learning letters relative to opportunity would be a warning sign, as would information, perhaps from standardized tests, of working memory deficiencies. Constraints on associative learning are best addressed, it is reasonable to assume, by more abundant exposure to letters and words. This advice is confirmed by purely behavioral data, for example from Reitsma (1983), who showed that children identified as reading disabled required more exposures to novel words to fix them in memory than other children did. While the suggestion of extended practice is hardly revolutionary, it does contrast to a degree with the dominant approach to reading difficulties, namely an emphasis on phonological awareness and explicit teaching of letter–sound relations.

We can draw other lessons from findings of genetically influenced variation in foundational processes supporting reading. Working memory will be a pivotal process at many points other than left-to-right decoding of words early in reading development. Assembling meaning from continuous text will also call upon working memory. The research implicating poor working memory as a factor affecting reading development is quite substantial (Fletcher, Lyon, Fuchs, & Barnes, 2018; Hulme & Snowling, 2009). Indeed, a variety of executive functions, working memory, inhi-

bition, processing speed, and naming speed, is substantially related to reading skills, and most of that covariation is driven by genetic influence (Christopher et al., 2016). This in turn suggests that instructional practices that load executive functions like working memory unnecessarily will be detrimental to reading-disabled children; such practices could range from asking them to decode overly long words to text construction that delivers information in sentences with unnecessarily complex syntactic structures such as deep embedding. Educators designing material, from elementary readers to standard textbooks, would do well to keep this type of genetically influenced constraint in mind as they do their work. Otherwise, children who labor under these constraints will not only miss out on needed successful encounters with print and information from texts, but run the risk of declining motivation to read as they become dispirited by their attempts to do so (Byrne, 2005).

These, then, are examples of how genetically informative research can help guide practice, and why educational professionals are likely to benefit from its continuation and dissemination. Researchers, too, have a role to play here, namely in ensuring that the products of their efforts are made readily available to the communities that can turn them into good practice.

9.9 Goals of Education

The available evidence tells us that in educational environments that are relatively uniform in terms of child attendance, curriculum, teacher quality, and so on, genetic endowment will be a restraining factor for some children (and of course an enhancing one for others). But the evidence also tells us that well-designed instructional interventions and continued support can help moderate individual risk (see Chap. 8), and in fact produce encouraging results at the whole-of-district level (Sadoski & Wilson, 2006). So an important question becomes, how much academic equality among students should an educational jurisdiction strive for with the use of strenuous instructional strategies to compensate individuals and groups?

The question is a real one because there are only so many hours in the school day and an intensive focus on one school subject, literacy say, must come at the expense of another. The problem is compounded by the fact that multivariate behavior-genetic analyses have shown that there is a considerable amount of pleiotropy among school subjects; children who are genetically challenged in literacy, for example, are often challenged for at least some of the same (genetic) reasons in mathematics (Willcutt et al., 2013). Students who are burdened with genes that act in this pleiotropic fashion, that is across both domains, will need extra remedial help in both, placing that much more time pressure on them and on the school system to provide such remediation.

Realistically, again, schools probably need to differentially value academic subjects in terms of where the extraordinary efforts are directed. Literacy would normally appear high on that list because of its importance for educational progress (and life) in general, with mathematics close behind because of the grounding it furnishes for the physical and social sciences, and of course for everyday life in a technological

society. These are not easy choices, but at least recognition of genetic limitations does force educators to make them. In contrast, clinging to the view that (almost) everything is down to the environment and that, therefore, there is a wide range of resources to intervene (improving teacher quality, reducing class sizes, ensuring greater involvement of families in the educational process, controlling time spent on social networking sites, ridding the environment of toxins, improving sleep habits and diet, and so on) may give rise to false hopes. Indeed, meta-analyses of reading interventions show that environmental factors (e.g., group size, number of intervention hours, SES), play a smaller role in intervention outcomes than we would intuitively expect (see Chap. 8 for a brief review). It is not that these factors are unimportant, but it is too optimistic by far to believe that these measures or a combination of them will level the playing field for students.

Thus, genetic endowment will remain an influence on achievement in actual educational systems, though educators committed to the notion that very high achievement is within the grasp of any child in one domain or another can take comfort in the fact that the environment can remain a substantial player in achievement. They can also take comfort in the long tradition of evidence that sustained practice over substantial periods of time can produce high-level skills in music, sporting activities, specialist academic pursuits, and other domains (Ericsson & Ward, 2007).

9.10 A Reiteration: Behavior-Genetic Studies Are About Variances, Not Means

In a thought experiment, imagine that a novel way of teaching reading was discovered that when introduced to schools raised the average skill level of all children to the degree that none now read at the low levels that previously would classify them as "dyslexic." None would now be hampered in the other parts of their education. It is a reasonable assumption that differences in, say, speed of reading, would remain and that they would be in part driven by genetic differences, but in practical terms, these would be of no more consequence than differences in speed of normal walking are now. Drawing out the implications of these observations, research should continue into the best way to teach literacy even though genes might continue to drive student differences. Findings of even high heritability for a trait do not imply that population averages cannot change, as indeed they have over time for the highly heritable variable of human height or, closer to home for the concerns of this article, for the almost equally heritable variable of intelligence (Dickens & Flynn, 2001).

9.11 Conclusions

- Genetic variability among students impacts individual differences in literacy. In most research conducted within a genetically informative framework to this point, genes account for half or more of the variance. But the quantitative estimates derived from each study need to be conditioned by important considerations, most notably the educational context in which it was undertaken. Environments that are both intensive and uniform will generate higher estimates of heritability than those that are less intensive and more variable. It would be quite inappropriate to import estimates from contexts like those in Scenario 1 to Scenario 2 situations. Doing so could weaken any motivation to change the educational circumstances that could be threatening the academic progress of Scenario 2 students.
- Quantitative estimates of heritability and environmentality can be affected by the way the calculations are done, as in our example of how results from multisite environments are standardized and combined.
- Heritability is about variability, not average levels of achievement. Comparisons across school districts, states, and countries are appropriate measures of averages. Further, even high levels of genetic influence should not inhibit the search for better ways of building and delivering the curriculum, nor of attending to any adverse academic circumstances of subgroups within a population.
- High heritability is not a recipe for inaction. Genes are never the full story. It is impossible to determine for any individual the degree to which genes versus environmental factors are at play. Given evidence that prevention and intervention efforts can ameliorate the adverse effects of genetic endowment, sustained efforts to remediate remain appropriate when children or groups are falling behind.
- Quantitative estimates do not, of themselves, tell us anything about how genes and aspects of the environment influence student achievement. Well-designed studies will build in measures other than literacy itself in an attempt to identify those factors. This is where hypotheses gained in other research can be of high value, such as when processes identified in purely phenotypic studies of reading development and dyslexia are incorporated into the research. The days of simply quantifying heritability and environmentality of literacy are, or ought to be, over.
- Despite the qualifications and cautions we have outlined, genetic endowment is a real player in relative levels of academic achievement, and will continue to be. Recognizing that fact can provide some relief for parents and teachers who may have been called on to bear more than their fair share of responsibility for children who struggle with one subject or another. But the same relief may be accompanied by greater pessimism about remediation, and the educational community needs to be proactive by developing ways to forestall unnecessarily gloomy attitudes about prognosis.
- The data we have surveyed also pose questions for educational systems about the extent they wish to compensate, at the individual or system level, for constraints imposed by genes. These are not easy questions to answer, but facing them squarely is a start.

Acknowledgements This research was supported by the Australian Research Council (DP0663498, DP0770805), the National Institute for Child Health and Human Development (HD27802, HD38526), the Swedish Research Council (2011–1905), and the Swedish Council for Working Life and Social Research (2011–0177).

References

Asbury, K., & Plomin, R. (2013). *G is for genes: The impact of genes on education and achievement*. Oxford, UK: Wiley-Blackwell.

Australian Education Union. (2012). *Devolution and education: A research report prepared by the Australian Education Union.* Retrieved from http://www.aeufederal.org.au/Publications/2012/Devandeducation.pdf.

Barnes, J. C., Wright, J. P., Boutwell, B. B., Schwartz, J. A., Connolly, E. J., Nedelec, J. L., et al. (2014). Demonstrating the validity of twin research in criminology. *Criminology, 52,* 588–626.

Blachman, B. A., Schatschneider, C., Fletcher, J. M., Murray, M. S., & Munger, K. (2014). Intensive reading remediation in grade 2 or 3: Are there effects a decade later? *Journal of Educational Psychology, 106,* 46–57.

Byrne, B. (2005). Theories of learning to read. In M. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 104–119). Oxford: Blackwell.

Byrne, B., Khlentzos, D., Olson, R. K., & Samuelsson, S. (2010). Evolutionary and genetic perspectives on educational attainment. In K. Littleton, C. Wood, & J. K. Staarman (Eds.), *International handbook of psychology in education* (pp. 3–33). Bingley, UK: Emerald Press.

Byrne, B., Olson, R. K., & Samuelsson, S. (2013). Subsample standardization in twin studies of academic achievement. In *Annual conference of the behavior genetics association*, July, Marseilles, France.

Byrne, B., Wadsworth, S., Boehme, K., Talk, A. C., Coventry, W. L., Olson, R. K., … Corley, R. (2013). Multivariate genetic analysis of learning and early reading development. *Scientific Studies of Reading, 17*, 224–242.

Carey, N. (2012). *The epigenetics revolution*. London, UK: Icon Books.

Chabris, C. F., Lee, J. J., Cesarini, D., Benjamin, D. J., & Laibson, D. I. (2015). The fourth law of behavior genetics. *Current Directions in Psychological Science, 24,* 304–312.

Christopher, M. E., Keenan, J. M., Hulslander, J., DeFries, J. C., Miyake, A., Wadsworth, S. J., et al. (2016). The genetic and environmental etiologies of the relations between cognitive skills and components of reading ability. *Journal of Experimental Psychology: General, 145,* 451–466.

Dar-Nimrod, I., & Heine, S. J. (2011). Genetic essentialism: On the deceptive determinism of DNA. *Psychological Bulletin, 137,* 800–818.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Educational Policy Analysis Archives, 8*. Retrieved from http://epaa.asu.edu/epaa/v8n1/.

Dickens, W. T., & Flynn, (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review, 108*, 346–369.

Dong, C., Taylor, M. P., Kristensen, L. J., & Zahran, S. (2015). Environmental contamination in an Australian mining community and potential influences on early childhood health and behavioural outcomes. *Environmental Pollution, 207,* 345–356.

Elbro, C., & Peterson, D. K. (2004). Long-term effects of phoneme awareness and letter sound training: An intervention study with children at risk for dyslexia. *Journal of Educational Psychology, 96,* 660–670.

Ericsson, K. A., & Ward, P. (2007). Capturing the naturally occurring superior performance of experts in the laboratory: Toward a science of expert and exceptional performance. *Current Directions in Psychological Science, 16,* 346–350.

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2018). *Learning disabilities: From identification to intervention* (2nd ed.). New York: Guilford.

Florestal, K., & Cooper, R. (1997). Decentralization of education: Legal issues. *Directions in Development Series*. Washington, DC: The World Bank. Retrieved from http://siteresources.worldbank.org/EDUCATION/Resources/278200-1099079877269/547664-1099080000281/Decent_ed_legal_issues_EN97.pdf.

Grasby, K. L., Coventry, W. L., Byrne, B., Olson, R. K., & Medland, S. E. (2016). Genetic and environmental influences on literacy and numeracy performance in Australian school children in grades 3, 5. 7, and 9. *Behavior Genetics, 46,* 627–648.

Grigorenko, E. L. (2007). How can genomics inform education? *Mind, Brain, and Education, 1,* 20–27.

Harlaar, N., Spinath, F. M., Dale, P. S., & Plomin, R. (2005). Genetic influences on early word recognition abilities and disabilities: A study of 7-year-old twins. *Journal of Child Psychology and Psychiatry, 46,* 373–384.

Haslam, N., & Kvaale, E. P. (2015). Biogenetic explanations of mental disorder: The mixed-blessings model. *Current Directions in Psychological Science, 24,* 399–404.

Hulme, C., & Snowling, M. J. (2009). *Developmental disorders of language and cognition*. Malden, MA: Wiley-Blackwell.

Kovas, Y., Haworth, C. M. A., Dale, P. S., & Plomin, R. (2007). The genetic and environmental origins of learning abilities and disabilities in the early school years. *Monographs of the Society for Research in Child Development, 72,* 1–160.

Kvaale, E. P., Haslam, N., & Gottdeiner, W. H. (2013). The "side effects" of medicalization: A meta-analytic review of how biogenetic explanations affect stigma. *Clinical Psychology Review, 33,* 782–794.

Little, C. W., Haughbrook, R., & Hart, S. A. (2017). Cross-study differences in the etiology of reading comprehension: A meta-analytical review of twin studies. *Behavior Genetics, 47,* 52–76.

Maher, B. (2008). Personal genomes: The case of the missing heritability. *Nature, 456,* 16–21.

Marope, M. (2010). The education system in Swaziland: Training and skills development for shared growth and competitiveness. *World Bank Working Paper No 188, Africa Region Human Development Department*, Washington, DC. Retrieved from http://siteresources.worldbank.org/EDUCATION/Resources/278200-1099079877269/education_system_in_Swaziland.pdf.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2002). *Mx: Statistical modeling* (6th ed.). VCU Box 900126, Richmond, VA 23298: Department of Psychiatry.

Ojanen, E., Ronimus, M., Ahonen, T., Chansa-Kabali, T., February, P., Jere-Folotiya, J., et al. (2015). GraphoGame—A catalyst for multi-level promotion of literacy in diverse contexts. *Frontiers in Psychology., 6,* 671.

Olson, R. K., Keenan, J. M., Byrne, B., & Samuelsson, S. (2014). Why do children differ in their reading and related skills? *Scientific Studies of Reading, 18,* 38–54.

Olson, R. K., Keenan, J. M., Byrne, B., Samuelsson, S., Coventry, W. L., Corley, R., … Hulslander, J. (2011). Genetic and environmental influences on vocabulary and reading development. *Scientific Studies of Reading, 15*, 26–46. https://doi.org/10.1080/10888438.2011.536128.

Petrill, S. A., Deater-Deckard, K., Thompson, L. A., Schatschneider, C., DeThorne, L. S., & Vandenbergh, D. J. (2007). Longitudinal genetic analysis of early reading: The western reserve reading project. *Reading and Writing, 20,* 127–146.

Poelmans, G., Buitelaar, J. K., Pauls, D. K., & Franke, B. (2011). A theoretical molecular network for dyslexia: Integrating available genetic findings. *Molecular Psychiatry, 16,* 365–382.

Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2013). *Behavioral genetics* (6th ed.). New York, NY: Worth Publishers.

Purdie, N., & Buckley, S. (2010). School attendance and retention of Indigenous Australian students. *Issues paper No 1, Closing the Gap Clearinghouse*. Australian Institute of Health and Welfare, and Australian Institute of Family Studies. Retrieved from http://www.aihw.gov.au/closingthegap/documents/issues_papers/ctg-ip01.pdf.

Reitsma, P. (1983). Printed word learning in beginning readers. *Journal of Experimental Child Psychology, 75,* 321–339.

Sadoski, M., & Wilson, V. L. (2006). Effects of a theoretically based large-scale reading intervention in a multicultural urban school district. *American Educational Research Journal, 43,* 137–154.

Samuelsson, S., Byrne, B., Hulslander, J., & Olson, R. (2009). Behavior-genetic analyses of literacy development in more and less transparent orthographies: A comparison of Scandinavian and U.S./Australian children. In *Annual conference of the society for the scientific study of reading*, July: Boston, MA, USA.

Samuelsson, S., Byrne, B., Olson, R. K., Hulslander, J., Wadsworth, S., Corley, R, … DeFries, J. C. (2008). Response to early literacy instruction in the United States, Australia, and Scandinavia: A behavioral-genetic analysis. *Learning and Individual Differences, 1*8, 289–295 (PMID: 19122888).

Snowling, M. J., & Hulme, C. (2005). *The science of reading: A handbook*. Oxford, UK: Blackwell.

Taylor, J., & Schatschneider, C. (2010). Genetic influence on literacy constructs in kindergarten and first grade: Evidence from a diverse twin sample. *Behavior Genetics, 40,* 591–602.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of word reading efficiency*. Austin, TX: Pro-ed.

Willcutt, E. G., Petrill, S. A., Wu, S., Boada, R., DeFries, J. C., Olson, R. K., et al. (2013). Comorbidity between reading disability and math disability: concurrent psychopathology, functional impairment, and neuropsychological functioning. *Journal of Learning Disabilities, 46,* 500–516.