Genetics and the Science of Reading

Stephanie Estrera[1,2], Hope Sparks Lancaster[3], and Sara A. Hart[1,2]

[1]Department of Psychology, Florida State University

[2]Florida Center for Reading Research, Florida State University

[3]Center for Childhood Deafness Language and Learning, Boys Town National Research Hospital

Paper should be cited as:

Estrera, S., Lancaster, H. S., and Hart, S. A. (2023). Genetics and the Science of Reading. *The Reading League Journal 4*(2). 18-25.

Author Note

Stephanie Estrera https://orcid.org/0000-0003-3340-1006
Hope Lancaster https://orcid.org/0000-0001-9924-9508
Sara A. Hart https://orcid.org/0000-0001-9793-0420

This work is supported by Eunice Kennedy Shriver National Institute of Child Health & Human Development Grant HD052120 and by the Institute of Education Sciences FIREFLIES Doctoral Training Fellowship through Grant R305B200020. Views expressed herein are those of the authors and have neither been reviewed nor approved by the granting agencies.

Correspondence concerning this article should be addressed to Stephanie Estrera, Florida State University, 1007 W. Call Street, Tallahassee, FL, 32306, United States. Email: sestrera@fsu.edu.

As is clear to many teachers, students in the same classroom differ widely in their reading skills. Students may differ in reading skills due to their previous reading instruction, their access to books, or their families' levels of access to economic resources. Students who have access to similar resources, or even attend the same school, can still differ greatly in their reading skills. One reason for this is that students each have their own genetic predispositions which function within the environments they're in to influence how they read.

Reading skill development is a result of the interplay between genetic (nature) and environmental (nurture) influences. Understanding the role of each of these on reading skills is valuable in creating and improving upon educational interventions and instructional practices. Understanding the nature and nurture of reading skills is done in the fi eld of behavioral genetics. Here we provide a review of core findings from the fi eld of behavioral genetics on reading science, followed by an overview of the key findings from modern behavioral genetics studies on the interplay of genetics and environments on reading skills. We will also discuss aspects of the work that remain unanswered and recommendations for what the field should research next to best serve teachers and students.

First, we must speak to the common misconception that genes predetermine a student's future educational outcomes—as if a student is destined to either be skilled at or struggle with reading. Reading skills are heritable, but they are also shaped by the student's environment. (See our previous paper in The Reading League Journal that presented a thorough explanation of heritability, van Dijk et al., 2022). Students need explicit and systematic reading instruction and a learning environment that ensures they reach their maximum potential, no matter their genetic predisposition.

Core Findings: Why Do Students Vary in Their Reading Skills?

There is over a century of evidence from behavioral genetics establishing that reading skills run through families. These studies tracked the prevalence of reading skills, including difficulties within families, and found that if someone has a parent who struggles with reading, they are more likely to struggle with reading (Snowling & Melby-Lervåg, 2016). However, these family studies do not differentiate whether this is due to families sharing genes or the same environments (e.g., living in the same home, having the same economic resources). One method that behavioral geneticists have used for almost a century to separate genetic from environmental effects is twin studies. Twin studies allow researchers to disentangle nature from nurture by comparing the differences in reading skills between identical and fraternal twins. This provides a way to estimate how much of students' differences in their reading skills are explained by genes and environments—the latter of which is further separated into two components: the shared environment (e.g., going to the same school, having the same number of books at home) and the nonshared environment (e.g., having different friends or different illnesses).

Core research, built on decades of scientific research, has established that genes and both components of the environment—the shared environment and the nonshared environment—matter for reading skills. The differences in reading skills are represented numerically by their genetic, shared environmental, and nonshared environmental estimates. Each estimate ranges from 0% to 100% and denotes how much of the individual differences in reading skills are associated with genes or the environment. A paper that summarized the results from 37 twin and family studies focused on reading comprehension found an average genetic estimate of 59%, an average shared environmental estimate of 15%, and an average nonshared environmental estimate of 29% (Little et al., 2017b). Therefore, even in the same classroom with the same instruction, we can expect that students will differ in their reading skills.

Behavioral geneticists have been interested in understanding the genetic and environmental influences on various reading skills, as well as how genes and environments influence the association between reading skills. What this research has shown is that there is evidence that the genes for one reading skill (and reading related skills) overlap with the genes for other reading skills—a finding that Plomin and Kovas (2005) coined the "generalist genes hypothesis." Rather than one specific set of genes being responsible for a specific reading skill, the generalist genes hypothesis is the idea that the same genes are responsible for why people are different from each other across different reading skills. For example, the genes responsible for one reading skill, such as decoding, would be the same as those responsible for other types of reading skills, like reading comprehension (Hart et al., 2010). Both practitioners and researchers benefit from understanding the overlap between the genes for various reading skills, because it provides a foundation of evidence that there is a shared genetic undercurrent to all reading processes and shows us just how linked reading skills are.

A fundamental research question within the field of reading is whether students with a reading disability (e.g., non-responders to explicit and systematic instruction) are distinctly different from skilled readers, or if reading disability simply represents the lower end of the spectrum of reading skills. Generally, behavioral genetics findings support the idea that the same genes that influence reading difficulties also influence typical reading skills, suggesting that reading disability is not a categorically distinct genetic disorder (Little & Hart, 2022). For example, Erbeli and colleagues (2018b) defined reading disability as having a combination of indicators of reading difficulty (sometimes called the hybrid model) and found that, like average reading skills, genetic and environmental influences matter for reading disability. Other behavioral genetics studies (e.g., Haworth et al., 2009) also found that reading disability was influenced by genetic and environmental factors.

Difficulties in reading also tend to highly co-occur with difficulties in other areas, such as language (e.g., developmental language disorder), attentional behaviors (e.g., attention deficit/hyperactivity disorder), and math (e.g., math learning disability; Paracchini, 2022). The high co-occurrence between these difficulties implies that developmental learning difficulties are not independent from one another. Behavioral genetic studies have illustrated genetic overlap between traits associated with developmental learning disorders, suggesting that the same genes underlie difficulties across domains (Daucourt et al., 2020; van Bergen et al., 2014).These findings show us that the genetic and environmental influences on reading skills do not differ between skilled and struggling readers. This work has contributed to our current understanding that reading disability represents the lower end of the spectrum of reading skills, and that spectrum of skills is influenced by an underlying spectrum of genetic and environmental influences. Therefore, there is no reading disability specific gene, or genes. Additionally, common genes underlie struggles across academically important skills, accounting for the common finding in science and schools that students who struggle in one area tend to struggle in others. This would suggest that we should not expect reading disability to be classified only in the absence of other struggles.

Key Findings from Modern Behavioral Genetics: Developmental Studies

In early elementary school, students build their reading skills through direct classroom instruction with a focus on decoding and comprehending the meaning of texts. As they progress and become more fluent, the emphasis shifts to content-area instruction, though there is opportunity within the classroom for teachers to refine their students' reading skills through content-area reading instruction (Capin et al., 2021). Just as it is obvious to teachers that students are different from each other in their reading skills at any given time, it is also obvious that some students go from beginning readers to fluent text comprehenders at different rates. It is possible to examine the genetic and environmental influences on that building of or change in reading skills over time.

When examining the genetic and environmental influences on reading skills measured as snapshots of different ages, we generally find that the higher the genetic estimates are, the older the students in the sample are (Daucourt et al., 2020; Little et al., 2017b). Rather than snapshots, it is possible to follow the same students and repeatedly measure reading skills as they get older. Longitudinal datasets are promising sources of insight because they provide a sharper image of how reading develops over time within a student. The genetic influences on a student's developmental change over time tend to be lower than skills measured (i.e., the environment) as a single snapshot in time. This implies that the environment has a stronger influence on how reading skills develop over time than when reading skills are measured at any single point in time (Kievit et al., 2022). For example, a twin study in Ohio examined how word identification, phoneme awareness, rapid naming, and pseudoword decoding in kindergarten changed across three school years (Petrill et al., 2010). The researchers found that genetic influences and shared environmental influences (e.g., going to the same school, having the same number of books at home) were important for kindergarten reading skills. Surprisingly, shared environmental influences strongly affected the change of reading skills. Almost 100% of the differences between how students changed on word identification and decoding were attributed to the shared environment. More recent studies using the same statistical methods have found similar results (Hart et al., 2013; Little et al., 2017a).

What is unknown about this work is why the environment matters more for change over time than when we measure students at any snapshot of time. Researchers need to figure out what it is in students' environments that is causing this effect. We hypothesize it's the direct instruction in reading they are receiving, but this has not been quantified yet.

Key Findings from Modern Behavioral Genetics: The Interconnection of Genes and the Environment

Since both genetic and environmental effects are of interest in understanding reading skill development, it is important to consider how they work together. Rather than understand nature versus nurture, modern behavioral genetics focuses on the interplay of nature and nurture. Genes express themselves through the environment, and there are methods to look beyond genetic and environmental independent effects by considering the ways readers' genes and their current environment work with one another (Harden, 2021). For example, the genes related to reading skills influence the amount students engage with text (Olson et al., 2017). Gene-environment interplay encapsulates the ways genetic and environmental influences depend on each other, which includes gene-environment correlations and gene-environment interactions.

Gene-Environment Correlations

Gene-environment correlations describe how exposure to certain environments are partially determined by genetics (Scarr & McCartney, 1983). There are three types of gene-environment correlations. Passive gene-environment correlations occur when the genes parents pass on to their children also influence the environment they create. For example, parents with strong reading skills are more likely to spend time reading and have a home library, meaning they pass along a rich home literacy environment and genes associated with higher reading skills. Evocative gene-environment correlation refers to the relation between an individual's genetic disposition on a trait, and the reaction others may have to that genetically driven trait. This is illustrated by a student who is genetically inclined to be a fl uent reader being more likely to be asked to read out loud in class, and therefore getting more practice reading. The third is active gene-environment correlation, which is when individuals seek out environments consistent with their own genetically influenced skills—like a student with a genetic predisposition toward struggling with reading demonstrating avoidant behaviors toward reading, resulting in less reading practice (van Bergen et al., 2014). Generally, the presence of gene-environment correlations reminds us that the environment students find themselves in is genetically influenced. We note with caution that gene-environment correlations provide a basis of understanding that genetically influenced traits may affect students' exposure to certain types of environments; however, a correlation does not imply causation.

Gene-Environment Interaction

Gene-environment interactions describe the ways in which one's environment strengthens or weakens the effect of one's genes on a trait, such as reading skills. Studies based in the United States have evidence of various school-based factors impacting the strength of the genetic influences on reading skills. For example, one twin study examining school quality and early reading skills found that students who attended higher quality schools had higher genetic influences on their reading skills than students at lower quality schools (Haughbrook et al., 2017). Another twin study analyzed the moderating influence of school socioeconomic status (SES) on third- and fourth-grade students' reading comprehension, finding that greater genetic variance was present in schools with lower SES (Hart et al., 2013). Alternatively, an Australian study that looked at students in Grades 3-9 examined moderating influences of home and school SES and found that genetic variation in reading skills was not influenced by SES (Grasby et al., 2016). The contrasting take home messages from these studies implies that the difference in environmental influences varies with context (i.e., the country and the educational structures one is operating in). This informs those of us who are interested in why students differ in reading skills by bringing us closer to understanding how, and for whom, certain environments impact reading skills.

What is clear from this work is that the reality of what influences an individual student in the classroom is extremely complex. As a field, behavioral genetics does not have the methods to fully understand why a given student is struggling. Given the complexity of collecting samples, such as twin samples, much of the scientific evidence from behavioral genetics is based on White, western, and well-resourced families. This means we don't yet have good evidence of the role of the full spectrum of environments that students are experiencing (Holden et al., 2022).

Key Findings from Modern Behavioral Genetics: Molecular Genetics

Technological advancements in more affordable and noninvasive measurements of human DNA have brought forth a new era of genetic research. If you have used 23andMe or other related ancestry commercial products, you have benefited from these new advancements. Outside of commercial use, molecular genetic technological advancements within the behavioral genetics fi eld can point us toward which genes are associated with reading skills. Initially, the findings from twin studies of moderate genetic effects on reading skills led researchers to try to find a specific "dyslexia gene." We have since learned that there is no such single gene, or even a handful of genes, that underlie reading skills (Paracchini, 2022). Instead, it is clear now that outside of a few specific case examples of families with unique genes associated with reading difficulties, for most students, hundreds to thousands of their genes contribute, each in a small way, to their reading skills.

A new type of study called genome wide association studies (GWASs) allows behavioral genetics researchers to identify the genes associated with reading skills. A GWAS identifies which and to what extent genetic variants (i.e., their letter, C, T, G, or A, formally called single nucleotide polymorphisms, or SNPs) are associated with a reading skill(s) at the population level. An even newer technique, called polygenic scores, allows us to add up all the genetic variants associated with reading skills, therefore quantifying the total genetic influence on reading skills into one number (Harden, 2021). Recently, a GWAS assessed the genetic associations with reading skills (e.g., word and nonword reading, phoneme awareness, spelling) in 19 different research samples, totaling over 33,000 people aged 5 to 26 (Eising et al., 2022). They found that genetic variants associated with reading skills were important for related cognitive skills like memory and educational attainment. The authors of a different GWAS for word reading found that the genes associated with word reading were also associated with educational attainment and neurodevelopmental disorders involved in learning and language-based difficulties (e.g., ADHD; Price et al., 2020). The largest GWAS of dyslexia to date included over 50,000 adults self-reporting a dyslexia diagnosis and identified 42 important genetic variants for dyslexia. The authors found that individuals with more genetic variants associated with dyslexia (i.e., higher polygenic scores) had lower reading and spelling accuracy, with a polygenic score of dyslexia explaining up to 6% of the differences between people in their reading traits (Doust et al., 2022). Altogether, this new and building research using molecular genetics techniques supports the idea that many genes of small effect influence reading skills;

these genes are not specific to reading skills but are in common with other traits such as language.

Being able to predict reading skills from an individual's genetics is still a very new research area, but it is thought that polygenic scores may be a valuable tool to identify children with reading disabilities early, potentially enabling learning support for children who may need more intervention even earlier (Shero et al., 2022). However, we cannot apply these methods in schools until the scientific community builds larger and more diverse samples of participants. The work reviewed above is from GWAS studies conducted mainly in populations with European ancestry. Polygenic scores that stem from GWASs of one population tend to be less predictive of traits in other populations. Thus, the continuation of large, international scientific collaborations may yield deeper insights in the future of the role of genetics on reading skills. Perhaps, in time, these approaches will grant us a way to reliably estimate the risk of reading disabilities so students with these disabilities are tended to earlier and with care.

Future Directions

Core and contemporary behavioral genetics research have clarified that genetic and environmental influences matter for students' reading skills. Our group is interested in thinking about how we can make behavioral genetics more applicable to teachers and students, so that the scientific findings can make a difference for students in classrooms. Our understanding of the genetics of reading skills is changing quickly, with advancements in twin research and molecular genetics. In the future, we may be able to use students' DNA to personalize education, but not yet. It is imperative that behavioral geneticists engage in team science, which includes educational stakeholders and shareholders, to ensure findings can lead to improvements in students' lives.

Acknowledgements

This work was supported by Eunice Kennedy Shriver National Institute of Child Health & Human Development Grants HD052120. Stephanie Estrera is supported by the Institute of Education Sciences FIREFLIES Doctoral Training Fellowship (R305B200020). Views expressed herein are those of the authors and have neither been reviewed nor approved by the granting agencies.

References

Capin, P., Stevens, E. A., Stewart, A. A., Swanson, E., & Vaughn, S. (2021). Examining vocabulary, reading comprehension, and content knowledge instruction during fourth grade social studies teaching. *Reading and Writing*, *34*(5), 1143–1170.

Daucourt, M. C., Erbeli, F., Little, C. W., Haughbrook, R., & Hart, S. A. (2020). A meta-analytical review of the genetic and environmental correlations between reading and attention-deficit/hyperactivity disorder symptoms and reading and math. *Scientific Studies of Reading, 24*(1), 23–56.

Doust, C., Fontanillas, P., Eising, E., Gordon, S. D., Wang, Z., Alagöz, G., Molz, B., 23andMe Research Team, Aslibekyan, S., Auton, A., Babalola, E., Bell, R. K., Bielenberg, J., Bryc, K., Bullis, E., Coker, D., Partida, G. C., Dhamija, D., Das, S., … Luciano, M. (2022). Discovery of 42 genome-wide significant loci associated with dyslexia. *Nature Genetics, 54*(11), 1621–1629.

Eising, E., Mirza-Schreiber, N., de Zeeuw, E. L., Wang, C. A., Truong, D. T., Allegrini, A. G., Shapland, C. Y., Zhu, G., Wigg, K. G., Gerritse, M. L., Molz, B., Alagöz, G., Gialluisi, A., Abbondanza, F., Rimfeld, K., van Donkelaar, M., Liao, Z., Jansen, P. R., Andlauer, T. F. M., … Fisher, S. E. (2022). Genome-wide analyses of individual differences in quantitatively assessed reading- and languagerelated skills in up to 34,000 people. *Proceedings of the National Academy of Sciences, 119*(35), e2202764119.

Erbeli, F., Hart, S. A., Wagner, R. K., & Taylor, J. (2018b). Examining the etiology of reading disability as conceptualized by the hybrid model. *Scientific Studies of Reading, 22*(2), 167–180.

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*(5), 627–648.

Harden, K. P. (2021). "Reports of my death were greatly exaggerated": Behavior genetics in the postgenomic era. *Annual Review of Psychology, 72*(1), 37–60.

Hart, S. A., Petrill, S. A., & Dush, C. M. K. (2010). Genetic influences on language, reading, and mathematics skills in a national sample: An analysis using the National Longitudinal Survey of Youth. *Language, Speech, and Hearing Services in Schools, 41*(1), 118–128.

Hart, S. A., Logan, J. A. R., Soden-Hensler, B., Kershaw, S., Taylor, J., & Schatschneider, C. (2013). Exploring how nature and nurture affect the development of reading: An analysis of the Florida Twin Project on Reading. *Developmental Psychology, 49*(10), 1971–1981.

Haworth, C. M. A., Kovas, Y., Harlaar, N., Hayiou-Thomas, M. E., Petrill, S. A., Dale, P. S., & Plomin, R. (2009). Generalist genes and learning disabilities: A multivariate genetic analysis of low performance in reading, mathematics, language and general cognitive ability in a sample of 8000 12-year-old twins. *Journal of Child Psychology and Psychiatry, 50*(10), 1318–1325.

Holden, L. R., Haughbrook, R., & Hart, S. A. (2022). Developmental behavioral genetics research on school achievement is missing vulnerable children, to our detriment. *New Directions for Child and Adolescent Development, 2022*(183–184), 47–55.

Kievit, R. A., Logan, J. A., & Hart, S. A. (2022). From the trajectory of heritability to the heritability of trajectories. *The Behavioral and Brain Sciences, 45*, e165.

Little, C. W., Hart, S. A., Quinn, J. M., Tucker-Drob, E. M., Taylor, J., & Schatschneider, C. (2017a). Exploring the co-development of reading fluency and reading comprehension: A twin study. *Child Development, 88*(3), 934–945.

Little, C. W., Haughbrook, R., & Hart, S. A. (2017b). Cross-study differences in the etiology of reading comprehension: A meta-analytical review of twin studies. *Behavior Genetics, 47*(1), 52–76.

Little, C. W., & Hart, S. A. (2022). Genetic and environmental influences on learning to read. In M Snowling, C. Hulme & K Nation (Eds), *The Science of Reading: A Handbook.* Cambridge, UK: Cambridge University Press.

Olson, R. K., Kennan, J. M., Byrne, B., & Samuelsson, S. (2017). Genetic and environmental influences on the development of reading and related skills. In K. Cain, D. L. Compton, & R. K. Parrila (Eds.), *Theories of Reading Development* (Vol. 15). John Benjamins Publishing Company.

Paracchini, S. (2022). The genetics of dyslexia. In M. Snowling, C. Hulme & K. Nation (Eds.), *The Science of Reading: A Handbook*. Cambridge University Press.

Petrill, S. A., Hart, S. A., Harlaar, N., Logan, J., Justice, L. M., Schatschneider, C., Thompson, L., DeThorne, L. S., Deater-Deckard, K., & Cutting, L. (2010). Genetic and environmental influences on the growth of early reading skills: Reading growth. *Journal of Child Psychology and Psychiatry, 51*(6), 660–667.

Plomin, R., & Kovas, Y. (2005). Generalist genes and learning disabilities. *Psychological Bulletin, 131*(4), 592–617.

Price, K. M., Wigg, K. G., Feng, Y., Blokland, K., Wilkinson, M., He, G., Kerr, E. N., Carter, T., Guger, S. L., Lovett, M. W., Strug, L. J., & Barr, C. L. (2020). Genome-wide association study of word reading: Overlap with risk genes for neurodevelopmental disorders. *Genes, Brain and Behavior, 19*(6).

Scarr, S. & McCartney, K. How people make their own environments: A theory of genotype-environment effects. *Child Development,* 54, 424-435 (1983).

Shero, J., van Dijk, W., Edwards, A., Schatschneider, C., Solari, E. J., & Hart, S. A. (2021). The practical utility of genetic screening in school settings. *Npj Science of Learning, 6*(1), 12.

Snowling, M. J., & Melby-Lervåg, M. (2016). Oral language deficits in familial dyslexia: A meta-analysis and review. *Psychological Bulletin, 142*(5), 498-545

van Bergen, E., van der Leij, A., & de Jong, P. F. (2014). The intergenerational multiple deficit model and the case of dyslexia. *Frontiers in Human Neuroscience*, 8.

van Dijk, W., Daucourt, M. C., & Hart, S. A. (2022). Understanding heritability in the context of reading ability and instruction. *The Reading League Journal, 3*(1), 24–34.