Country-by-genotype-by-environment interaction in childhood academic achievement

Eveline L. de Zeeuw[a] and Dorret I. Boomsma[a,1]

Why children differ in their cognitive abilities, including their academic achievement, is a function of their environment and genotype, and this function may be complex, involving gene–environment (GxE) interaction and gene–environment (GE) covariation. The paper by Figlio et al. (1) in PNAS on academic achievement in reading and mathematics considers the importance of GxE interaction in a unique statewide sample of twin children from Florida. GxE interaction can be thought of as the genetic control over sensitivity to environmental conditions or, alternatively, as the environment determining the expression of the genotype. Several scenarios may be distinguished: Across different environments, the same genes are expressed, but their absolute or relative influences vary (quantitative differences), or different genes are expressed across different environments (qualitative differences). Obviously, a combination of these two scenarios is also possible. A third alternative states that under GxE interaction, environmental influences vary conditional on a child's genotypic value (e.g., a low or high genotypic value for mathematical ability) (2).

As to how GxE interaction may manifest, several theoretical models each offer their own prediction. The diathesis-stress model predicts that genetic vulnerability, or diathesis, increases the likelihood of a trait in the presence of environmental stress. It also predicts that the heritability of the trait will be higher for children in risk environments (3). In contrast, the bioecological model predicts that risk environments will mask genetic differences between children and enriched environments will amplify genetic differences (4). In studies of human behavior, researchers usually rely on cross-sectional designs with groups of related or genotyped individuals exposed to different environments to assess quantitative and qualitative GxE interactions.

Such a cross-sectional design was employed by Figlio et al. (1): Twin children's academic achievement in mathematics and reading in primary school was observed under different socioeconomic status (SES) circumstances. Birth and public school records were matched for all children in Florida from birth cohorts from 1994–2002, resulting in a large representation of twins from disadvantaged homes and a broad range of social and economic circumstances. The study explicitly tested the bioecological hypothesis, also known as the Scarr–Row interaction (5), of increased genetic and decreased environmental variation (i.e., higher heritability) in children from advantaged backgrounds, but found no evidence to support this bioecological hypothesis.

The results of the study by Figlio et al. (1) critically depended on the analysis of reading and mathematics data from twins and, as a comparison group, siblings. The classical twin design estimates resemblance in twins as a function of their zygosity. If a trait is influenced by genes, the resemblance of "identical" or monozygotic (MZ) twin pairs will be larger than that of fraternal or dizygotic (DZ) twin pairs, who share, on average, 50% of their segregating genes, just like non-twin siblings (6). Information on zygosity is not available from administrative databases, and differences in twin resemblance in the Florida study were assessed in same-sex (SS) and opposite-sex (OS) (i.e., male-female, female-male) twin pairs. The last group is always DZ, while the first group will be a mixture of MZ and DZ twin pairs, and the number of OS twin pairs is expected to be roughly equal to the number of DZ SS twin pairs.

A critical assumption in the paper is that the correlation between SS and OS DZ twin pairs is equal and that responsible genes do not differ in their expression between boys and girls. We will address this assumption by comparing resemblances for reading and mathematics achievement from a large group of twin pairs from the Netherlands Twin Register for whom zygosity was known (7). Correlations for OS and SS twin pairs and, as a critical evaluation, also the SS correlations separately for the MZ and DZ twin pairs are presented in Fig. 1. We clearly see that correlations for all SS twin pairs are in between the correlations for MZ and DZ twin pairs and that the correlations for the DZ twin pairs are very similar to the correlations for the

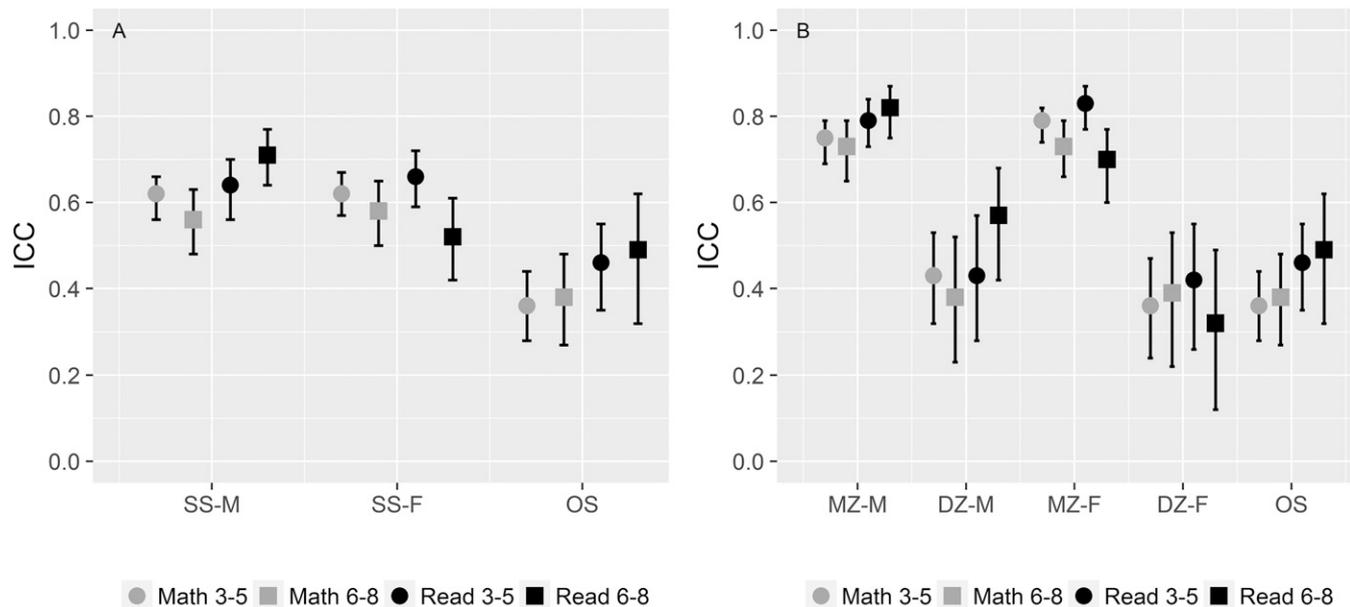www.pnas.org/cgi/doi/10.1073/pnas.1718938115

Fig. 1. Intraclass correlations (ICCs) and their 95% confidence intervals for mathematics and reading achievement in Dutch twin pairs across the same grades (3–8) as reported by Figlio et al. (1). ICCs for same-sex (SS) male (SS-M), SS female (SS-F), and opposite-sex (OS) twin pairs (*A*) and ICCs for monozygotic (MZ) male (MZ-M), dizygotic (DZ) male (DZ-M), MZ female (MZ-F), DZ female (DZ-F), and OS twin pairs (*B*) are given. For each ICC, the number of twin pairs is between 100 and 300.

OS twin pairs, demonstrating that the assumption of the method used by Figlio et al. (1) indeed holds.

In Fig. 1, the difference in resemblance between the SS and OS twin pairs is noticeably similar in the twin pairs from The Netherlands compared with the twin pairs from Florida, and the correlations in Dutch DZ twin pairs are close to the Florida sibling correlations, again validating the approach that Figlio et al. (1) took. Here, we also note that earlier heritability estimates for cognitive ability without information on zygosity were comparable to estimates obtained from classical twin studies (8) and to estimates from population-based studies where multiple family relationships between children were available but zygosity was unknown (9).

The lack of evidence for a bioecological GxE interaction is an important finding, as it is in contrast to a large meta-analysis of genotype × SES interaction on academic achievement and intelligence concluding that the effect existed in the United States, but not in European and Australian countries, and that it was even reversed in The Netherlands (10). As stated by Figlio et al. (table 1 of ref. 1), in the data from Florida twin pairs, the GxE interaction also was uniformly signed in the reversed direction, although not always significant. Their findings thus not only provide no evidence for the bioecological model but even point to the opposite (i.e., the diathesis-stress model), as the genetic variance was significantly smaller in children from a higher SES background for several of the academic achievement measures. On the other hand, the environmental variance was significantly larger in children from a higher SES background for most measures. As may be seen in the study by Figlio et al. (SI appendix of ref. 1), this pattern of lower genetic and higher environmental variance under high SES circumstances was consistent across all alternative SES measures and proportions of MZ twin pairs in the SS group. This implies that their results are more comparable to what has been found for GxE interaction in The Netherlands than to the previous US findings.

One explanation for the dissimilar interaction results in the earlier studies conducted in the United States and other Western countries concerned the differences in socioeconomic inequality.

Variation in social and economic circumstances is larger in the United States, with higher income inequality (11), child poverty rates (12), and differences in educational opportunities (13). Figlio et al. (1) clearly failed to replicate earlier US studies in their large US sample, indicating either that GxE interaction US studies have been underpowered (false-positive results) or that even within a country, the phenomenon may not be universal and may depend on state.

In the study by Figlio et al. (1), maternal education is the main index for environmental (dis)advantage. There is substantial resemblance in academic achievement between parents and offspring, which is mainly due to genetic inheritance (e.g., ref. 14). Thus, the reason for an association between SES and academic achievement, as again confirmed by the strong association between maternal education and children's mathematics and reading scores (table 2 of ref. 1), is likely genetically mediated, representing a form of GE covariation (i.e., children with "favorable" genotypes tend to grow up in favorable environments). To take such GE covariation into account, the statistical GxE interaction model used by Figlio et al. (1) included a main effect of SES on academic measures; as such, the genotype × SES interaction effect only moderated the variance in academic achievement that is independent from SES (14, 15).

Figlio et al. (1) discuss exciting options for future research, when insight into GxE interaction might be enhanced with molecular genetic data. The idea is that if SES interacts with academic achievement, then the influence of genetic variants that have been established as causal genetic variants for academic achievement will differ across socioeconomic backgrounds. Figlio et al. (1) write that it may take some time before sample sizes of genome-wide association (GWA) studies are large enough to test this hypothesis at the measured genotype level. This would be the case when limiting to robustly associated genetic variants, so-called genome-wide associated hits, but, nowadays, polygenic scores (PGSs), which include the information from all genetic variants, already provide a realistic research opportunity. Sample sizes of the GWA studies on which the PGSs are based have increased

tremendously, resulting in a growing predictive power of PGSs. The PGSs based on the most recent GWA study for academic attainment explained over 3% of differences between individuals (16, 17). A GxE interaction in PGS analyses would be reflected in a difference in magnitude of explained variance between socioeconomic backgrounds. For the bioecological model of GxE, the PGS is expected to explain more variance in academic achievement in children from high SES backgrounds. A recent study saw no evidence of an interaction between PGS and family SES on educational achievement or on general cognitive ability (17).

GxE interaction studies on academic achievement have not only focused on the home environment but also on aspects of the school environment. For example, heritability of academic achievement did not depend on education type (e.g., Montessori) (18), but the effect of genes on reading ability differed across levels of teacher quality (19). GxE interaction is complex, and it is important to gather evidence from multiple environments, samples, and research designs, especially when a next step in educational research involves developing personalized education policies to enhance children's performance at school. Knowledge on the existence of genotype × SES interaction for academic achievement is important as it means that by targeting children's social and economic circumstances, the realization of their genetic potential can be facilitated irrespective of family background.

1 Figlio DN, Freese J, Karbownik K, Roth J (2017) Socioeconomic status and genetic influences on cognitive development. *Proc Natl Acad Sci USA* 114:13441–13446.
2 Schwabe I, Boomsma DI, van den Berg SM (2017) Increased environmental sensitivity in high mathematics performance. *Learn Individ Differ* 54:196–201.
3 Rende R, Plomin R (1992) Diathesis-stress models of psychopathology: A quantitative genetic perspective. *Appl Prev Psychol* 1:177–182.
4 Bronfenbrenner U, Ceci SJ (1994) Nature-nurture reconceptualized in developmental perspective: A bioecological model. *Psychol Rev* 101:568–586.
5 Scarr-Salapatek S (1971) Race, social class, and IQ. *Science* 174:1285–1295.
6 Boomsma D, Busjahn A, Peltonen L (2002) Classical twin studies and beyond. *Nat Rev Genet* 3:872–882.
7 de Zeeuw EL, van Beijsterveldt CEM, Glasner TJ, de Geus EJC, Boomsma DI (2016) Arithmetic, reading and writing performance has a strong genetic component: A study in primary school children. *Learn Individ Differ* 47:156–166.
8 Benyamin B, Wilson V, Whalley LJ, Visscher PM, Deary IJ (2005) Large, consistent estimates of the heritability of cognitive ability in two entire populations of 11-year-old twins from Scottish mental surveys of 1932 and 1947. *Behav Genet* 35:525–534.
9 Schwabe I, Janss L, van den Berg SM (2017) Can we validate the results of twin studies? A census-based study on the heritability of educational achievement. *Front Genet* 8:160.
10 Tucker-Drob EM, Bates TC (2016) Large cross-national differences in gene × socioeconomic status interaction on intelligence. *Psychol Sci* 27:138–149.
11 OECD (2011) *Divided We Stand: Why Inequality Keeps Rising* (OECD Publishing, Paris).
12 Chapple S, Richardson D (2009) *Doing Better for Children* (OECD Publishing, Paris).
13 Hauser RM (1970) Educational stratification in the United States. *Sociol Inq* 40:102–129.
14 Swagerman SC, et al. (2017) Genetic transmission of reading ability. *Brain Lang* 172:3–8.
15 Purcell S (2002) Variance components models for gene-environment interaction in twin analysis. *Twin Res* 5:554–571.
16 Okbay A, et al.; LifeLines Cohort Study (2016) Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* 533:539–542.
17 Selzam S, et al. (2017) Predicting educational achievement from DNA. *Mol Psychiatry* 22:267–272.
18 Schwabe I, Boomsma DI, Zeeuw EL, Berg SM (2016) A new approach to handle missing covariate data in twin research: With an application to educational achievement data. *Behav Genet* 46:583–595.
19 Taylor J, Roehrig AD, Soden Hensler B, Connor CM, Schatschneider C (2010) Teacher quality moderates the genetic effects on early reading. *Science* 328:512–514.