

BIROn - Birkbeck Institutional Research Online

Austerberry, Chloe and Fearon, P. and Ronald, Angelica and Leve, L.D. and Ganiban, J.M. and Natsuaki, M.N. and Shaw, D.S. and Neiderhiser, J.M. and Reiss, D. (2022) Early manifestations of intellectual performance: evidence that genetic effects on later academic test performance are mediated through verbal ability in early childhood. *Child Development* 93 (2), e188-e206. ISSN 0009-3920.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/46034/>

Usage Guidelines:

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html>

or alternatively

contact lib-eprints@bbk.ac.uk.

**Early Manifestations of Intellectual Performance: Evidence That Genetic Effects on
Later Academic Test Performance are Mediated Through Verbal Performance in Early
Childhood**

Abstract

Intellectual performance is highly heritable and robustly predicts lifelong health and success but the earliest manifestations of genetic effects on this asset are not well understood. This study examined whether early executive function (EF) or verbal performance mediate genetic influences on subsequent intellectual performance, in a US-based sample of 561 adoptees (57% male) and their birth and adoptive parents (70% and 92% White, 13% and 4% African American, 7% and 2% Latinx, respectively). Genetic influences on children's academic performance at 7 years were mediated by verbal performance at 4.5 years ($\beta = .22$, 95% CI [0.08, 0.35], $p = .002$) and not via EF, indicating that verbal performance is an early manifestation of genetic propensity for intellectual performance.

Keywords. Adoption, Executive Function, Verbal Ability, Academic Achievement, Development.

Early Manifestations of Intellectual Performance: Evidence That Genetic Effects on Later Academic Test Performance are Mediated Through Verbal Performance in Early Childhood

Intelligence and academic performance are among the strongest predictors of lifelong success, health and longevity (Deary, Weiss, & Batty, 2010; Hummer & Hernandez, 2013; Kosik et al., 2018), and both are substantially heritable (Bouchard & McGue, 1981; Kovas et al., 2013). However, the evidence base is small regarding the earliest manifestations of genetic influences on intellectual performance. These manifestations are important because they may be in the causal chain from genetic influences to later intellectual performance and because they are also likely to have an indirect influence on intellectual development through interaction with caregiving and learning environments. We use a parent-offspring adoption design to examine two likely candidates for this early manifestation: Executive function (EF) and verbal performance in early to middle childhood. Our results are the first to document whether early EF or verbal performance have a mediating role, linking genetic influences on later intellectual performance in middle childhood and possibly also in adulthood. By identifying which of these, EF or verbal performance, serve as a principal manifestation of genetic influences, our results pave the way for investigations into how children's interactions with parents and teachers from early childhood onwards amplify or diminish these favorable outcomes.

Intelligence and academic performance are powerful predictors of psychological wellbeing, health, longevity, years of education, income and employment status (Deary et al., 2010; Hummer & Hernandez, 2013; Kosik et al., 2018). Conversely, lower intellectual performance is associated with all-cause mortality and clinically important increases in the severity of psychopathology (Deary et al., 2010; Kosik et al., 2018; Yew & O'Kearney, 2013). Furthermore, there is evidence to suggest that academic performance in adolescence

may have a negative causal connection with internalizing and externalizing problems in emerging adulthood (Wolchik, Tein, Sandler, & Kim, 2016). Consequently, promotion of intellectual performance in childhood may have broad effects across development, including improving educational, occupational and health outcomes, and diminishing the likelihood of some psychiatric problems. As a result, research aimed at understanding the processes involved in the early development of intellectual performance is crucial and may help uncover mechanisms that can be modified, not only to promote intellectual development, but also to promote a wide range of positive life outcomes and reduce the risk of psychopathology.

Intelligence and academic test performance have been reported to be highly heritable, especially as children get older, rising from 20–60% in childhood and adolescence to 50–80% in adulthood (Bouchard & McGue, 1981; Haworth et al., 2010; Kovas et al., 2013). Consequently, some have argued that environmental factors must play only a minor role in intellectual development (Plomin, 2018). However, twin and adoption studies provide evidence that environmental factors can have notable main effects and moderating effects on intellectual outcomes (Capron & Duyme, 1989; Kendler, Turkheimer, Ohlsson, Sundquist, & Sundquist, 2015; Neiss & Rowe, 2000; Tucker-Drob & Bates, 2016). There is also evidence from the recent surge of literature using measured genotypes to examine genetic nurture (Bates et al., 2018; Kong et al., 2018; Wertz et al., 2020)—including studies that have combined polygenic scores with the adoption design (Cheesman et al., 2020; Domingue & Fletcher, 2020)—suggesting that parents influence children’s educational outcomes not only through direct genetic transmission but also through environmentally mediated pathways. An additional, unheralded, mechanism is that the environment may have an amplifying effect on genetic influences, through evocative gene-environment correlation (r_{GE}). This occurs when an individual’s genetically influenced characteristics systematically evoke responses from

their environment that, in turn, enhance or ‘canalize’ genetic influences (Scarr & McCartney, 1983). As these evoked environmental conditions correlate with genetic influences their influence could be entirely masked by estimates of genetic main effects. Dickens and Flynn (2001) explore in detail the possibility that this process of amplification operates in the context of cognitive abilities across generations to account for rising levels of intelligence in successive cohorts of children and adults. While there is some evidence from phenotypic, twin and polygenic score research of evocative rGE in infant and early childhood cognitive development (Lugo-Gil & Tamis-LeMonda, 2008; Tucker-Drob & Harden, 2012; Wertz et al., 2020), the evidence base is small and the Dickens and Flynn hypothesis has never been robustly tested across the span of development within a generation. For these environmental amplification effects to be examined in detail, it is important to know at which developmental periods they may exert their influence on intellectual outcomes. For influences occurring early in development it is crucial to identify the earliest manifestations of genetic advantage because these are likely the features that elicit the favorable environmental responses that amplify genetic effects.

In spite of the great importance of identifying the early manifestations of genetic influences on lifespan intellectual performance, the evidence base is small with regards to what these early manifestations might be. There is some indication that childhood scholastic performance from 6–7 years old onwards may be an early indicator of genetic advantage for intellectual performance in adulthood. For example, higher genome-wide polygenic scores of total years of education achieved by adulthood (EA PGS) predict stronger reading and math test performance at 6, 7, 12 and 16 years (Allegrini et al., 2019; Armstrong-Carter et al., 2020; Belsky et al., 2016; Selzam et al., 2017). This is supported by evidence that adopted children’s math and reading performance at age 7 years is partially predicted by their birth parents’ reading and math test performance (Borriello et al., 2020; Cioffi et al., 2021). These

associations are not confounded by direct caregiving effects because adopted children and birth parents share genes, but birth parents do not provide the postnatal rearing environment. Furthermore, as the birth parent outcomes were measured in adulthood, the observed phenotypic associations between birth parents and children are akin to an “instant longitudinal study” from childhood to adulthood (Plomin, 1986) because, although these studies do not include longitudinal data from childhood to adulthood, they identify genetic factors accounting for the association between academic test performance in childhood and intellectual performance of biological relatives in adulthood. It remains less well understood whether there are earlier markers of genetic effects on lifespan intellectual performance than academic test performance from age 6–7 years onwards. There is mixed evidence from one longitudinal study (the Dunedin Study): Children in the sample with higher EA PGS began talking earlier, based on parent ratings of developmental milestones at 3 years old, but did not score any better in the Peabody Picture Vocabulary Test at 3 years old (Belsky et al., 2016). However, from 5 years old and onwards, children in the study with higher EA PGS scored higher on tests of intelligence. Additionally, there is some evidence from adoption studies that birth parent intelligence in adulthood predicts adopted children’s executive function (EF), verbal and non-verbal intelligence from 1–3 years old (Leve, DeGarmo, et al., 2013; Plomin, Fulker, Corley, & DeFries, 1997), indicating that these early abilities may be markers of genetic effects on adult intelligence. This is consistent with evidence, firstly, that EF and verbal performance in early childhood are partially heritable, including from as early as two years old, at which point the heritability of both are fairly low—around 20%—(Gagne, Asherson, & Saudino, 2020; Galsworthy, Dionne, Dale, & Plomin, 2000) and throughout early and middle childhood, by which point the heritability of EF and verbal performance appears to be approximately 60% (Davis, Haworth, & Plomin, 2009a; Polderman et al., 2007). Secondly, these findings are consistent with evidence that early

childhood EF and verbal performance predict subsequent intellectual performance (Duncan et al., 2007; McClelland, Acock, Piccinin, Rhea, & Stallings, 2013; Yu, McCoach, Gottfried, & Gottfried, 2018). For example, there is longitudinal evidence from six studies that reading, verbal performance and attention at school entry robustly predict later school math and reading test performance (Duncan et al., 2007). It seems likely, based on these converging bodies of research, that EF and verbal performance in early and middle childhood are early manifestations of genetic effects on later intellectual performance. However, no research has used an adoption design to combine these streams of evidence and investigate whether early—and apparently heritable—EF and verbal performance mediate genetic influences on later intellectual outcomes.

Early manifestations of genetic effects on intellectual performance are important to understand, firstly, because they may directly influence the development of later intellectual performance. Secondly, because they likely have an indirect influence on intellectual development through interaction with caregiving and learning environments that plausibly sustain and amplify these early manifestations. However, as there is limited evidence of what the early manifestations are and precisely when they manifest, research is not yet in a position to rigorously explore hypotheses about evocative effects of genetic influences underlying intellectual development on caregiving and learning environments. A critical first step is to identify the very early expressions of genetic advantage in intellectual performance using a longitudinal parent-offspring adoption study.

We examine the confirmatory hypothesis that genetic effects linked to adult intellectual performance have their impact on child reading and math test performance at 7 years of age through two early-appearing pathways: via EF and verbal performance from 27 months to 6 years. This hypothesis is addressed in two steps. First, we examine at what age, or ages—between the ages of 27 months to 6 years—is there evidence of genetic effects on

EF and verbal performance. Second, if there appear to be genetic effects on early EF and verbal performance, we test for mediation of genetic effects on reading and math test performance at 7 years old via each of these pathways. We employ birth mother general intellectual performance—captured using a latent composite of intelligence, reading and math test performance—as a proxy for genetic influences. As adopted children and their birth mothers share genes but birth mothers do not provide the postnatal rearing environment, the adoption design eliminates the influence of birth mothers on the postnatal environment. Phenotypic associations between adopted children and their birth mothers will thus be taken to imply genetic effects. However, correlations between birth mothers and their adopted offspring can represent a combination of genetic and intrauterine effects. Birth fathers, who play an equal role to mothers in contributing to the child's genotype, provide an estimate of genetic effects that is not confounded by intrapartum effects. Consequently, we use a smaller subsample of birth fathers for replications of the birth mother analyses. Although birth parents tend to correlate on measures of intelligence (Bouchard & McGue, 1981) and birth father replications can only be considered quasi-independent rather than fully independent replications, broadly speaking, they provide convergent evidence regarding genetic—as opposed to intrauterine—effects on children's intellectual performance.

Method

Participants

Participants were drawn from the Early Growth and Development Study (EGDS), a US-based, longitudinal, prospective adoption study of 561 linked sets of adopted children and their birth mothers ($n = 554$), birth fathers ($n = 210$) and adoptive parents (562 adoptive fathers and 569 adoptive mothers; numbers do not sum to 561 adoptive mothers and fathers because the sample includes 41 same-sex parent families and 15 additional adoptive parents who entered the family after the original couple adopted the child) (Leve et al., 2019; Leve,

Neiderhiser, et al., 2013). EGDS data were collected in two cohorts, recruited through 45 adoption agencies in 15 states across the US (Leve et al., 2019): The first, a sample of 361 adopted children and their birth and adoptive families and, the second, a sample of 200. While some of the variables to be used in the proposed analysis were collected in both cohorts, others were only collected in one cohort. For a breakdown of the number of participants by each variable, see Table 1 and Figure 1. The variables used in the present analyses were collected in 2003–2013 (cohort I) and in 2007–2017 (cohort II).

Mean age of children at adoption placement was 5.58 days postpartum ($SD = 12.4$; median = 2; range = 0–91). Over half of the children were male (57%). Adoptive parents were typically White (adoptive parent 1: 92%; adoptive parent 2: 90%). The remainder were African American (adoptive parent 1: 4%; adoptive parent 2: 5%), Hispanic or Latinx (adoptive parent 1: 2%; adoptive parent 2: 2%), and “other” (adoptive parent 1: 2%; adoptive parent 2: 3%). At the time of adoption, adoptive parents were typically in their mid- to late-thirties (adoptive parent 1: $M = 37.4$, $SD = 5.6$; adoptive parent 2: $M = 38.3$, $SD = 5.8$), married or cohabiting (adoptive parent 1: 98%, adoptive parent 2: 100%), college educated and with a combined median income above \$100,000. Birth parents were typically White (mothers: 70%; fathers: 70%). The remainder were African American (mothers: 13%; fathers: 12%), Hispanic or Latinx (mothers: 7%; fathers: 10%), and “other” (mothers: 10%; fathers: 9%). At the time of adoption, birth parents were typically in their mid-twenties (mothers: $M = 24.4$, $SD = 6.0$; fathers: $M = 26.1$, $SD = 7.8$), married or cohabiting (mothers: 6.1%; fathers: 14.0%), had less than a college education, and median household incomes below \$25,000. There is no evidence of selective placement in EGDS (Leve et al., 2019). Additional information about the recruitment, composition and representativeness of the sample is reported elsewhere (Leve et al., 2019). EGDS assessments are ongoing and occurred in intervals of 9 months to 2 years. We use data collected from birth parents at 18

months, 4.5 and 7 years postpartum and data collected from adoptive parents and adoptees when the children were 27 months, 4.5 years, 6 years and 7 years old.

Ethics

Ethical approval was obtained from institutional review boards at [masked for review]. Informed consent was obtained from all adult participants ahead of research participation and assent was obtained from children at age 7 years.

Measures

Using structural equation models, incorporating confirmatory factor analysis (CFA), we created the latent variables (displayed in Figure 1) across each of the domains outlined below. Prior to hypothesis testing, we ran longitudinal measurement models, assessing the fit of individual domains across all timepoints. Model fit was good in all of these models, supporting the use of latent variables.

Birth Parent General Intellectual Performance

As displayed in Figure 1, we created a latent variable of birth parent general intellectual performance—with the indicators of intelligence, reading and math test performance listed below—as a proxy for genetic influences on children. Latent measurement drawing on a diverse range of indicators was justified by the internal consistency (birth mother $\alpha_R = .84$; birth father $\alpha_R = .85$) and bivariate correlations among measures of birth parent intelligence and academic test performance in the EGDS sample (Table 1), and the ‘generalist genes’ literature which reports that approximately a third of the genetic variance of reading and math performance is in common with general intelligence (g) (Davis, Haworth, & Plomin, 2009b; Plomin & Kovas, 2005).

Wechsler Adult Intelligence Scale-III. We administered the 28-item Information subtest (Wechsler, 1997) to birth parents at 18 months postpartum. It loads onto the verbal comprehension index of the full measure and is considered to be a representative measure of

g (g loading = .79) (Kaufman & Lichtenberger, 1999). We used standardized scores, based on age.

Woodcock-Johnson Tests of Achievement III. At 4.5 or 7 years postpartum, we administered birth parents four subtests: (1) 76-item Letter-Word Identification, measuring reading decoding; (2) 32-item Word Attack, capturing decoding and phonetic coding; (3) 98-item Reading Fluency, measuring reading speed and semantic processing speed; (4) 160-item Math Fluency, indexing math and numerical performance (Woodcock, McGrew, & Mather, 2001). We used T-scores, standardized to have a mean of 50 and standard deviation of 10.

Child Executive Function

As displayed in Figure 1, we created latent measures of executive function (EF) at 27 months ($r = .15$), 4.5 years ($\alpha_R = .38$) and 6 years ($\alpha_R = .41$), with and the EF indicators listed below. Use of latent variables was justified by the model fit (RMSEA = .03, SRMR = .05) and longitudinal stability of EF from 27 months to 4.5 years ($\beta = .74$, 95% CI [.41, 1.07], $p < .001$) and from 4.5 years to 6 years ($\beta = .70$, 95% CI [.45, .95], $p < .001$) in the longitudinal measurement model, as well as by the concurrent and longitudinal correlations between these EF measures, in the EGDS sample (Table 1) and wider literature (Montgomery & Koeltzow, 2010).

Stroop Task. At 27 months, we administered the fruits-animals Stroop, modified by the EGDS team based on Kochanska, Murray, and Harlan (2000). There were six trials, each scored on a scale from 1 to 3 (1 = incorrect on item and size; 2 = correct item, wrong size; 3 = correct item and size). The trials had strong internal consistency ($\alpha = .85$) and were averaged to form a scale score. At 4.5 and 6 years, we administered the 16-trial day-night Stroop (Gerstadt, Hong, & Diamond, 1994), which has robust construct validity and internal and test-retest reliability (Montgomery & Koeltzow, 2010). Each trial had one point for a

correct answer. Trials had strong internal consistency ($\alpha = .85$) and were summed, resulting in a score between 0 and 16.

Gift Delay Task. At 27 months, children participated in a gift delay task similar to the one described by Kochanska, Murray, Jacques, Koenig, and Vandegeest (1996). We coded the videotaped task based on how often the child (1) peeked, (2) touched the gift, and (3) used distraction strategies. In line with Leve, DeGarmo, et al. (2013) we averaged the 3 items to form a total score of inhibitory control, with higher scores indicating higher inhibition ($\alpha = .54$; $r = .08, .32, \text{ and } .46$ among items).

Guessing Game. At 4.5 years old, children completed a task adapted from the Goldsmith and Rothbart (1999) laboratory assessment of temperament (Lab TAB) to measure their inhibitory control when told not to turn around or peek at hidden toys. The task was coded from 1 (Not at all) to 5 (Continually) on: “How often did the child keep their back turned around when asked to?”.

Forbidden Gift. We measured inhibitory control in the 4.5-year-olds using a forbidden gift task modified from the Lab TAB, which was videotaped and coded from 1 (Very True) to 3 (Not true) on whether: “The child asked for the gift”.

Dinky Toys. This inhibitory control task modified from the Lab TAB involved the 4.5- and 6-year-olds being asked to comply with rules about how to interact with a box of toys. We rated the task on: “The degree to which the child follows or violates instructions” from 1 (violates rules) to 3 (follows all instructions).

Go-NoGo. At 6 years, we administered a Go-NoGo task (Nosek & Banaji, 2001). In this 84-trial version, trials were divided into two blocks, the first of which contained only Go trials (when the child should press a button) and the second an equal combination of Go trials and NoGo trials (in which children are expected to inhibit their prepotent response by

refraining from pressing a button). We measured selective attention and inhibition using the percentage of correct responses in the second block to both Go and NoGo stimuli.

Child Verbal Performance

We created latent variables at 27 months ($r = .74$), 4.5 years ($\alpha_R = .62$) and 6 years ($\alpha_R = .76$) with the indicators of verbal performance listed below. Our decision to use latent variables combining these indicators was guided by the model fit (RMSEA = 0.06; SRMR = 0.07) and longitudinal stability from 27 months to 4.5 years ($\beta = .47$, 95% CI [.35, .59], $p < .001$) and from 4.5 years to 6 years ($\beta = .76$, 95% CI [.66, .87], $p < .001$) in the longitudinal measurement model, as well as by the concurrent and bivariate correlations between these measures in the EGDS sample (Table 1) and wider literature (Sim, Thompson, Marryat, Ramparsad, & Wilson, 2019).

Language Development scale (LDS). Adoptive parents separately completed a measure of child language development at 27 months, based on the number of words that the child is reported to use spontaneously from a list of 310 items (Achenbach & Rescorla, 2000). Reports from both parents were used as indicators in the verbal performance latent variable at 27 months. Using nationally standardized normed scores, we converted raw scores to percentiles that reflected the child's language performance relative to same-age peers (Achenbach & Rescorla, 2000). LDS scores have moderate to high correlations ($r = .66 - .87$) with scores on standardized vocabulary tests (Klee et al., 1998; Rescorla & Alley, 2001) and are reported to have the best predictive validity performance of the language screening tools (Sim et al., 2019).

Test of Preschool Early Literacy (TOPEL). We administered three subscales to 4.5-year-olds (Lonigan, Wagner, & Torgesen, 2007): (1) 36-item Print Knowledge, measuring knowledge of the alphabet, written language conventions and written form; (2) 35-item Definitional Vocabulary, assessing definitional and single-word oral vocabulary; (3) 27-item

Phonological Awareness, measuring word elision and blending. We used standard scores, derived from the distribution of the raw scores. The TOPEL has high internal consistency ($\alpha = .86 - .96$) and test-retest reliability ($r = .81 - .89$), moderate predictive validity ($r = .40 - .62$) and moderate to high concurrent validity ($r = .59 - .77$) (Lonigan, Allan, & Lerner, 2011).

Wechsler Preschool and Primary Scale of Intelligence III (WPPSI). We administered the vocabulary assessment to 6-year-olds, measuring learning, comprehension and verbal expression of vocabulary (Wechsler, 2002). Raw scores from the 50-item measure were converted to standardized scores from 1 to 19, based on the responder's age.

Dynamic Indicators of Basic Early Literacy Skills (DIBELS). We administered four sets of procedures and assessments to 6-year-olds: (1) 16-item Initial Sound Fluency (ISF), measuring phonemic awareness; (2) Letter Naming Fluency (LNF), capturing proficiency in naming upper- and lower-case letters, using a list of 110 letters; (3) Phoneme Segmentation Fluency (PSF), assessing proficiency in fluently segmenting three- and four-phoneme words into their individual phonemes, using a list of 24 words; (4) Nonsense Word Fluency (NWF), testing understanding of the alphabetic principle, including letter-sound correspondence, using a list of 50 nonsense words (Good & Kaminski, 2002). ISF and LNF have good test-retest reliability ($r = .88 - .93$) and robustly predict later reading performance (Kaminski & Good, 1996). Raw scores, which represent the number of items a child has answered correctly in one minute, were converted to percentiles, reflecting verbal performance relative to same grade level peers, based on nationally standardized normed scores (Good & Kaminski, 2002).

Child Academic Test Performance

Justified by the high genetic correlations between reading and math performance in childhood (Davis et al., 2009b; Plomin & Kovas, 2005), and the internal consistency ($\alpha = .87$)

and bivariate correlations in the EGDS sample (Table 1), we created a latent variable to estimate child academic test performance at 7 years old, drawing on the same four indicators of reading and math performance that were administered to birth parents from the Woodcock-Johnson Tests of Achievement III (Woodcock et al., 2001)—see Figure 1.

Covariates

We included adoption openness, sex of child, and prenatal risk as covariates. We used a mean standardized composite of birth mother and adoptive parent reported adoption openness, using a 4-item measure (Ge et al., 2008), averaged across ratings provided at 9, 18, and 27 months postpartum. We collected birth mother reports of maternal and pregnancy complications, labor and delivery complications and neonatal complications at 5 months postpartum and scored them based on the 76-item McNeil-Sjostrom Scale for Obstetric Complications (McNeil, Cantor-Graae, & Sjostrom, 1994). We used a weighted total prenatal risk score based on work by Marceau et al. (2016).

Data Analysis

We conducted our primary analyses using birth mother and child data only and used data from a smaller subsample of birth fathers to carry out a quasi-independent replication. Although the birth father sample is the largest ever recruited in a prospective parent-offspring adoption study, it has reduced statistical power compared to the birth mother analyses. Thus, we anticipated that comparisons between results for birth mothers and birth father would focus on the magnitude of the path coefficients rather than p values or confidence intervals.

We tested our hypothesis in two steps, in the lavaan package (Rosseel, 2012) in R 4.0.0, using structural equation modelling, which combines a measurement model (also known as CFA) with a structural model testing the proposed causal relations. First, we constructed longitudinal models examining: (1) Whether EF and verbal performance were stable across 27 months, 4.5 years and 6 years, and predicted academic test performance at 7

years; and (2) whether there were genetic effects on child EF, verbal performance and academic test performance. Second, if the models were consistent with mediation of genetic effects on academic test performance at 7 years through early EF or verbal performance, we ran mediation models examining the indirect effects of birth parent intellectual performance on child academic test performance at 7 years. We included the covariates in all of our models and we used bootstrapping with 5000 repetitions to test the indirect effect in the mediation models (Bollen & Stine, 1990). Based on recommendations by Hu and Bentler (1999), we used a combination rule, according to which model fit was considered adequate if $SRMR < .09$ and $RMSEA < .06$.

Variable sample sizes are reported in Table 1 and Figure 1. The primary source of missing data in child EF models using birth mother data was child EF measures at 4.5 years. In child verbal performance models using birth mother data it was the DIBELS ISF subscale. In birth father and child models, it was missing information on birth father intellectual performance. The data used in the analyses were not missing completely at random [Little's MCAR $\chi^2(4598) = 4884.36, p < .01$]. We ran an attrition analysis using the Missing Value Analysis function in SPSS, which creates an indicator variable identifying variables that contain missing values. This indicator value is then used to compare group means among different variables in the dataset, using the t-test procedure. The attrition analysis revealed that the patterns of missingness for the majority (69%) of study variables were related to the observed values of one or more other variables in the dataset. Full results from the attrition analysis are available from the authors on request. This analysis ruled out the possibility that the data were missing completely at random (MCAR), which occurs when the probability of being missing is the same for all cases and there is no systematic association between the missingness of the data and any other values, observed or missing. It was not possible to rule out the possibility that the data were missing not at random (MNAR), which is when the

missingness of the data is systematically related to unobserved data. However, the associations found in the attrition analysis are consistent with the data being missing at random (MAR), which occurs when the missingness of a variable is systematically related to the observed but not the unobserved data. Full information maximum likelihood (FIML) and multiple imputation (MI) are both suitable for data that is MAR, and are of comparable performance (Allison, 2003). We prioritized FIML for the results that we report and, additionally, re-ran the models using MI, with 100 imputations. Overall, FIML and MI produced equivalent results. The few discrepancies between them are reported in the supplementary materials. Full results from the models using MI are available from the authors on request.

Sensitivity Analyses

We conducted sensitivity analyses to examine: (1) the impact of the removal of earlier time points on associations between birth parent general intellectual performance and child EF, verbal performance and academic test performance; (2) whether the age at which birth parents were administered measures of intellectual performance was associated with their intellectual performance and, if so, whether birth parent age confounded the associations between birth parents and children; and (3) if the indirect association between birth parent intellectual performance and academic test performance at 7 years, via children's earlier verbal performance still held when the mediation models were re-computed using only the math subscale of the academic test performance measure at 7 years. The third sensitivity analysis was conducted as a robustness check to rule out the likelihood that mediated effects on academic test performance via verbal performance were simply due to the content overlap between the measures of early verbal performance and the reading subscales of the academic performance outcome measure at 7 years old. By way of comparison, the EF mediation

models were also re-computed, using only the math subscale as the outcome, rather than the latent measure of academic test performance.

Results

Means, standard deviations, sample sizes and bivariate correlations between study variables are presented in Table 1.

Early Executive Function and Later Academic Test Performance

Birth Mother Effects

As shown in Figure 2, birth mother intellectual performance was directly associated with EF at 27 months ($\beta = .33$, 95% CI [.11, .23], $p = .004$) and academic test performance at 7 years ($\beta = .25$, 95% CI [.10, .40], $p = .001$), and the total effect on academic test performance at 7 years was: $\beta = .31$, 95% CI [.19, .43], $p < .001$. There was no direct effect ($\beta = -.13$, 95% CI [-.51, .25], $p = .507$) or total effect ($\beta = .14$, 95% CI [-.06, .34], $p = .158$) of birth mother intellectual performance on EF at 4.5 years. Nor was there a significant direct effect ($\beta = .10$, 95% CI [-.15, .34], $p = .450$) or total effect ($\beta = .21$, 95% CI [-.02, .44], $p = .076$) at 6 years. The model accounted for 45% of the variance in EF at 27 months, 49% of the variance in EF at 4.5 years, 79% of the variance in EF at 6 years and 20% of the variance in academic test performance at 7 years. The sensitivity analysis revealed that when the 27-month timepoint was dropped from the model, effects of birth mother intellectual performance did not carry forward to 4.5 years (Figure S1a, Supplementary Material). Nor did they carry forward to 6 years, when 27 months and 4.5 years were removed from the model (Figure S1b, Supplementary Material).

In the mediation model examining whether the effect of birth mother intellectual performance on academic test performance at 7 years was mediated via EF at 27 months, the total effect of birth mother intellectual performance on academic test performance at 7 years was statistically significant ($\beta = .31$, 95% CI [.19, .43], $p < .001$). The indirect effect,

mediated through EF at 27 months, was 32% of the total effect and not statistically significant ($\beta = .10$, 95% CI $[-.29, .49]$, $p = .614$). Model fit: $\chi^2(68) = 209.19$, $p < .001$, CFI = .91, RMSEA = .07, SRMR = .06.

Effects on Math Performance. As in the original model that was being re-computed, in the sensitivity analysis re-computing the mediation analysis with the latent academic test performance variable at 7 years old replaced with the math fluency subscale of the Woodcock-Johnson, the indirect effect of birth mother intellectual performance on math test performance at 7 years old, mediated through child executive function at 27 months, was small and not statistically significant ($\beta = .05$, 95% CI $[-.29, .39]$, $p = .754$). The indirect effect was 36% of the total effect and half the size (50%) of the indirect effect in the original model. Model fit: $\chi^2(52) = 160.94$, $p < .001$, CFI = .83, RMSEA = .09, SRMR = .06.

Birth Father Effects

As in the birth mother model, birth father intellectual performance directly predicted academic test performance at 7 years ($\beta = .27$, 95% CI $[.05, .50]$, $p = .017$; Figure 3) and the total effect was: $\beta = .34$, 95% CI $[.16, .52]$, $p < .001$. The effect estimate of birth father intellectual performance on EF at 27 months was numerically similar to the effect estimate in the birth mother model, although it was not statistically significant in the birth father model ($\beta = .27$, 95% CI $[-.89, .63]$, $p = .141$). As in the birth mother model, birth father intellectual performance did not directly predict EF at 4.5 years ($\beta = .16$, 95% CI $[-.27, .59]$, $p = .474$), although (unlike in the birth mother model) the total effect reached statistical significance ($\beta = .36$, 95% CI $[.07, .65]$, $p < .016$). At 6 years, neither the direct effect ($\beta = -.03$, 95% CI $[-.43, .37]$, $p = .883$) or the total effect ($\beta = .26$, 95% CI $[-.04, .56]$, $p = .094$) reached statistical significance, and the effect estimates were similar to those in the birth mother model. The birth father model accounted for 43% of the variance in EF at 27 months, 55% of

the variance in EF at 4.5 years, 73% of the variance in EF at 6 years and 19% of the variance in academic test performance at 7 years.

As in the birth mother mediation model, the total effect of birth father intellectual performance on academic test performance at 7 years was statistically significant ($\beta = .31$, 95% CI [.13, .49], $p = .001$). The indirect effect of birth father intellectual performance on child academic test performance at 7 years, mediated through child EF at 27 months, was 22% of the total effect, not statistically significant ($\beta = .07$, 95% CI [-.22, .35], $p = .640$) and of comparable (small) magnitude to the birth mother results. Model fit: $\chi^2(68) = 120.12$, $p < .001$, CFI = .95, RMSEA = .04, SRMR = .07.

Effects on Math Performance. As in the original model that was being re-computed, in the sensitivity analysis re-computing the mediation analysis using the math fluency subscale of the Woodcock-Johnson at 7 years old (rather than the latent measure of academic test performance), the indirect effect of birth father intellectual performance on math test performance at 7 years old, mediated through child executive function at 27 months, was small and not statistically significant ($\beta = .02$, 95% CI [-.13, .18], $p = .768$). The indirect effect was 5% of the total effect and 29% the size of the indirect effect in the original model. Model fit: $\chi^2(36) = 49.98$, $p = .061$, CFI = .96, RMSEA = .03, SRMR = .07.

Early Verbal Performance and Later Academic Test Performance

Birth Mother Effects

As displayed in Figure 4, birth mother intellectual performance directly predicted child verbal performance at 4.5 years ($\beta = .35$, 95% CI [.21, .49], $p < .001$) and the total effect at 4.5 years was: $\beta = .38$, 95% CI [.24, .53], $p < .001$. Birth mother intellectual performance also directly predicted child academic test performance at 7 years ($\beta = .13$, 95% CI [.02, .25], $p = .026$) and the total effect at 7 years was: $\beta = .32$, 95% CI [.20, .44], $p < .001$. There was no evidence of direct effects of birth mother intellectual performance on verbal

performance at 27 months ($\beta = .09$, 95% CI $[-.05, .22]$, $p = .201$) or 6 years ($\beta = -.02$, 95% CI $[-.18, .13]$, $p = .768$), although there was a significant total effect at 6 years ($\beta = .28$, 95% CI $[.14, .43]$, $p < .001$). The model accounted for 3% of the variance in verbal performance at 27 months, 32% of the variance in verbal performance at 4.5 years, 62% of the variance in verbal performance at 6 years and 51% of the variance in academic test performance at 7 years. A sensitivity analysis revealed that when verbal performance at 27 months and 4.5 years were dropped from the model, effects of birth mother intellectual performance carried forward to verbal performance at 6 years (Figure S2, Supplementary Material).

In the mediation model, examining whether the effect of birth mother intellectual performance on academic test performance at 7 years was mediated via verbal performance at 4.5 years, the total effect of birth mother intellectual performance on academic test performance at 7 years was statistically significant ($\beta = .32$, 95% CI $[.20, .44]$, $p < .001$). The direct effect of birth mother intellectual performance on child academic test performance at 7 years was not statistically significant ($\beta = .10$, 95% CI $[-.04, .24]$, $p = .141$) and the indirect effect, mediated through child verbal performance at 4.5 years, was statistically significant ($\beta = .22$, 95% CI $[.08, .35]$, $p = .002$) and 68% of the total effect. Model fit: $\chi^2(81) = 211.74$, $p < .001$, CFI = .92, RMSEA = .06, SRMR = .06.

Effects on Math Performance. In the sensitivity analysis that re-computed the mediation analysis with the latent academic test performance variable at 7 years old replaced with the math fluency subscale of the Woodcock-Johnson, the findings were similar to those in the original model that was being re-computed. As in the original model, the indirect effect of birth mother intellectual performance on math test performance at 7 years old, mediated through child verbal performance at 4.5 years, was statistically significant ($\beta = .14$, 95% CI $[.03, .24]$, $p = .011$). The indirect effect was 88% of the total effect and just under two thirds

the size (64%) of the indirect effect in the original model. Model fit: $\chi^2(46) = 149.43, p < .001$, CFI = .87, RMSEA = .07, SRMR = .06.

Birth Father Effects

The model presented in Figure 4 was replicated in a sub-sample of children and their birth fathers. The model did not converge when the data at 27 months were included, so this timepoint was dropped from the model. As in the birth mother model, birth father intellectual performance significantly predicted child verbal performance at 4.5 years ($\beta = .37$, 95% CI [.11, .62], $p = .005$)—see Figure 5. Similar to the birth mother findings, there was no evidence of direct effects of birth father intellectual performance on verbal performance at 6 years ($\beta = .08$, 95% CI [−.21, .38], $p = .575$) and the total effect was significant ($\beta = .36$, 95% CI [.14, .60], $p = .002$). Unlike in the birth mother model, there was no evidence of direct effects of birth father intellectual performance on child academic test performance at age 7 years ($\beta = .09$, 95% CI [−.13, .30], $p = .433$), although, as in the birth mother model, the total effect at 7 years was significant ($\beta = .33$, 95% CI [.14, .51], $p = .001$). The model accounted for 18% of the variance in child verbal performance at 4.5 years, 63% of the variance in verbal performance at 6 years and 50% of the variance in academic test performance at 7 years. A sensitivity analysis revealed that, as in the birth mother sample, when verbal performance at 4.5 years was removed from the model, effects of birth father intellectual performance carried forward to verbal performance at 6 years (Figure S3a, Supplementary Material). When verbal performance at 4.5 and 6 years were dropped from the model, the effect of birth father intellectual performance on academic test performance at 7 years became significant (Figure S3b, Supplementary Material).

Similar to the birth mother mediation model, the total effect of birth father intellectual performance on academic test performance at 7 years was statistically significant ($\beta = .32$, 95% CI [.13, .50], $p = .001$). The direct effect of birth father intellectual performance on child

academic test performance at 7 years old was not statistically significant ($\beta = .12$, 95% CI $[-.09, .32]$, $p = .254$) and the indirect effect, mediated through child verbal performance at 4.5 years was statistically significant ($\beta = .20$, 95% CI $[.04, .36]$, $p = .016$) and explained 63% of the total effect. The numerical estimates were similar to those in the birth mother model. Model fit: $\chi^2(81) = 132.20$, $p < .001$, CFI = .96, RMSEA = .04, SRMR = .07.

Effects on Math Performance. In the sensitivity analysis, re-computing the mediation model using only the math subscale at 7 years old, the effects of birth father intellectual performance continued to be mediated by verbal performance. As in the original model that was being re-computed, there was a significant indirect effect of birth father intellectual performance on child math performance at 7 years old, mediated via child verbal performance at 4.5 years ($\beta = .09$, 95% CI $[.02, .16]$, $p = .018$). The indirect effect was 20% of the total effect and 45% the size of the indirect effect in the original model. Model fit: $\chi^2(46) = 54.55$, $p = .181$, CFI = .98, RMSEA = .02, SRMR = .07.

Discussion

Results were consistent with our hypothesis that effects of genetic influences on academic test performance at 7 years old are mediated by children's early verbal performance. Birth mother and birth father general intellectual performance each predicted child verbal performance from 4.5 years onwards, but not at 27 months, and genetic effects on academic test performance at 7 years of age were mediated through verbal performance at 4.5 years. This is consistent with the large literature on genetic influences on children's verbal performance (Stromswold, 2001) and extends the evidence by suggesting that verbal performance from 4.5 years old is an early manifestation of genetic influences on later intellectual performance. As the birth parent outcomes were measured in adulthood, the associations between birth mother or father intellectual performance and child verbal performance at 4.5 years are akin to "instant longitudinal" associations (Plomin, 1986),

indicating that early verbal performance may be a marker of genetic effects, not only on academic test performance at 7 years, but also on general intellectual performance in adulthood.

The association between birth parent intellectual performance and child executive function (EF) at 27 months, previously reported by Leve, DeGarmo, et al. (2013), was limited to this single occasion of measurement and did not reliably carry forward to 4.5 or 6 years in either the birth mother or birth father models. Additionally, the EF mediation models did not provide evidence of mediation of effects on academic test performance at 7 years through EF at 27 months.

Evidence that verbal performance from 4.5 years old may be an early manifestation of genetic influences on later intellectual performance converges with findings from the polygenic score literature. For instance our findings are in line with results from the Born in Bradford study, reporting that genome-wide polygenic scores of total years of education achieved by adulthood (EA PGS) predicted a composite measure of academic test performance (including aspects of verbal performance) in 6–7-year-old school children (Armstrong-Carter et al., 2020). However, our results provide evidence in a younger age group—preschool 4.5-year-old children. The absence of effects, in our sample, of birth parent intellectual performance on verbal performance at 27 months is at odds with detection in the Dunedin Study of a positive association between EA PGS and age of first words spoken, reported by parents when their children were 3 years old (Belsky et al., 2016). However, our findings are consistent with evidence from the same study (Dunedin) that, while there was no association between children’s EA PGS and their scores in the Peabody Picture Vocabulary Test at 3 years old, from 5 years onwards higher EA PGS predicted higher scores of intelligence (captured by composite measures of verbal and non-verbal performance). Our findings are consistent with evidence that in infancy individual differences in verbal

performance appear to be influenced to a greater degree by the shared environment than by genetic differences (Galsworthy et al., 2000) but that by middle childhood, heritability of verbal and non-verbal cognitive performance is higher and the shared environmental component reduces (Davis et al., 2009a). Our results are also in line with evidence that the cross-time correlations for genetic influences on cognitive outcomes are low in early childhood and increase substantially across childhood (Tucker-Drob & Briley, 2014), as well as with evidence that from middle childhood the same genetic influences on cognitive skills predominate, increasing in magnitude as children get older (Briley & Tucker-Drob, 2013). As noted by Briley and Tucker-Drob (2013), one possible explanation for higher heritability of verbal and non-verbal cognitive performance by the time children reach school age is that when children enter formal schooling, standardized educational practices somewhat equalize environmental differences between them, allowing genetic differences to have greater influence on individual differences. An additional explanation—which is compatible with our findings, as well as with the reviewed literature on the increasing heritability of cognitive performance throughout childhood and increasing stability of genetic influences as children age—is that transactional mechanisms of gene-environment interplay amplify genetic effects through processes such as evocative and active gene-environment correlation (Scarr & McCartney, 1983).

Limitations and Future Directions

It is unclear whether the inconsistency of EF effects reflects a lack of effects of birth parent intellectual performance on child EF at later timepoints, and the absence of mediation of genetic effects on intellectual performance via EF, or a failure to operationalize EF sufficiently reliably at these occasions of measurement. Although the EF measures used in the present study were less internally consistent than the measures of verbal performance, the use of latent variables corrected for attenuation by error and the temporal stability of the EF

latent variables was high. Compared to the temporal stability of verbal performance, the temporal stability of EF was higher from 27 months to 4.5 years and equivalent at 4.5 years to 6 years. It is also a possibility that EF was less predictive of later academic test performance than verbal performance due to high content overlap between indicators of verbal performance and the indicators of academic test performance that were included. However, this concern is somewhat mitigated by the results from the sensitivity analyses examining effects on only the math indicator of academic performance; the effects of birth parent intellectual performance continued to be mediated via verbal performance at 4.5 years old. This implies that verbal performance from 4.5 years is an early marker of genetic influences on a wider range of scholastic outcomes in middle childhood than simply those that are verbally oriented.

As our aim was to identify the earliest manifestations of genetic influences on later intellectual outcomes, it was important to include measures of EF and verbal performance from as early as 27 months in some of our analyses. However, as the 27 month measures miss important variance that is likely influenced by genetic pathways, estimates of effects on later child outcomes in the models that control for EF and verbal performance this early are substantially prone to omitted variable bias. Models not controlling for the earliest timepoint (which are thus less prone to this bias) are presented in the supplementary materials.

While our findings have the potential to aid the development of promotive and preventative interventions, they are unable to resolve uncertainty about whether early verbal performance is a liability-index (i.e. there are shared genetic factors that influence both verbal performance and subsequent academic test performance) or a causal mediator of genetic effects on subsequent academic test performance (i.e. limited verbal development would block the development of the skills necessary to perform well in academic tests) (Kendler & Neale, 2010). Each would have important but different implications for interventions in

childhood. Although both suggest that low verbal performance is a risk factor for low academic test performance, the latter suggests that early intervention targeted at verbal performance might offset risk, whereas the former might be an indication in favor of more sustained support. Future research should be aimed at testing these alternatives, through longitudinal examination of academic test performance following interventions directly on early verbal performance.

It is a strength of the current analysis that we controlled for the influence of the prenatal environment, by including a measure of prenatal risk and through replicating the analyses in the birth father sample. However, the lack of statistical power to accurately estimate the influence of birth father genetic effects is a limitation. Sufficiently powered research is needed on the influence of birth father contributions to intellectual outcomes. Birth father models are not fully independent replications and almost all of the measures of birth mother and birth father intelligence and academic test performance were correlated, suggesting the possibility of assortative mating, confounding, and partner interaction effects. In spite of the potential issues with spousal concordance, the birth father data add strength to our study—fathers play an equal role to mothers in contributing to the child’s genotype, provide a control for inpartum effects and are under-researched relative to mothers in developmental research. The role of birth fathers as a control for inpartum effects is somewhat threatened by the potential for fathers to have indirect effects on fetal development through, for example, contributing to the family dynamics in the home, stress level of the mother and material resources accessible to the mother. However, the likelihood of this confounding our results is diminished by the fact that the rates of birth parent cohabitation in the sample were low.

All behaviour genetics findings represent ‘what is’ in a particular sample and cultural context rather than what ‘could be’ in a different context (Plomin et al., 2017). Consequently,

it may be that there are features of the cultural milieu experienced by the US-based adopted children in our sample, that ‘transmit’ low-level genetic differences into differences in academic test performance to a greater or lesser degree than other cultural contexts might. Investigations into the representativeness of the EGDS sample have found that participating adoptive families appear to be representative of the US population (Leve, Neiderhiser, et al., 2013). However, relative to the birth parents, they are higher socioeconomic status (SES) (Leve et al., 2019), which may bias findings. It cannot be assumed that the conclusions of this study hold for children reared in low SES environments, particularly as SES appears to moderate genetic effects on intellectual outcomes (Capron & Duyme, 1989; Tucker-Drob & Bates, 2016). There is evidence from the UK Biobank that EA PGS are more predictive of educational outcomes among non-adopted than adopted children, and that children in the lowest decile of polygenic score for educational attainment reach a significantly higher level of education if they are adopted than if they are not adopted. (Cheesman et al., 2020). This converges with evidence from the US that children with low pre-adoption IQ scores experience substantial IQ score gains when adopted into high-SES families (Duyme, Dumaret, & Tomkiewicz, 1999), as well as with evidence that adoptees tend to academically out-perform their non-adopted biological siblings (Kendler et al., 2015). Collectively, these results indicate that genetic influences on education may be mediated by rearing environments or the wider cultural contexts that are associated with different rearing environments. Additionally, they suggest that estimates of direct genetic effects on academic outcomes may include mechanisms of gene-environment correlation and interaction, pointing to the possibility that genetic differences correlate and interact with different environmental mechanisms in different sociocultural contexts. There is evidence to suggest that different ethnic groups in the US and UK may exhibit different trajectories of verbal development (Saccuzzo, Johnson, & Russell, 1992; Zilanawala, Kelly, & Sacker, 2016). For example, in

the UK Millennium Cohort Study the ethnic groups in the sample had different odds of being in high or low performing profiles of verbal development in early childhood and these observed differences were mediated by the home learning environment, family routines, and the psychosocial environment (Zilanawala et al., 2016). Such findings illustrate the nuances of verbal development in different contexts and suggest that our results might not hold in samples from different cultural and ethnic groups or socio-economic circumstances, within or outside of the US. It remains unclear how mechanisms of gene-environment interplay influence the development of academic outcomes in a diverse range of cultural contexts. Most behaviour genetics research—including the present study—is conducted in developed countries and majority White samples. Replication of these methods in other countries and sociodemographic groups is needed and until then it cannot be assumed that the present findings generalize to other cultural contexts. Our interest in identifying a mediator in the association between birth parent and adopted offspring intellectual performance stems, in part, from an overarching aim to understand how rearing and learning environments may amplify the early manifestations of genetic influences on intellectual performance. However, it was not possible to form hypotheses about evocative effects of genetic influences underlying intellectual development without first identifying an early manifestation of genetic advantage that might elicit favorable and amplifying effects from parents. Now that we have identified early verbal performance as a likely mediator of genetic influences on lifespan intellectual outcomes, we can posit early caregiving and learning conditions that might amplify genetic advantage. Children's verbal performance predicts parenting quality—including dimensions of parenting such as, sensitivity, positive regard, cognitive stimulation, and responsiveness—which in turn predicts reading performance (Lugo-Gil & Tamis-LeMonda, 2008; Tucker-Drob & Harden, 2012). Consequently, future research should explore whether these aspects of parenting amplify genetic advantage in verbal performance.

Conclusion

This is the first study to examine whether early EF or verbal performance mediate genetic effects on later intellectual performance. Effects of birth parent intellectual performance on child academic test performance at 7 years old were mediated through verbal performance at 4.5-years-old but were not mediated by early EF. These findings suggest that early verbal performance may be a manifestation of genetic advantage for lifespan intellectual outcomes. Based on the importance of intellectual performance for lifelong health and adjustment, the apparent role of early verbal performance in intellectual development represents a critical finding.

References

- Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms & profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth & Families.
- Allegrini, A. G., Selzam, S., Rimfeld, K., von Stumm, S., Pingault, J. B., & Plomin, R. (2019). Genomic prediction of cognitive traits in childhood and adolescence. *Molecular Psychiatry, 24*, 819–827. doi:10.1038/s41380-019-0394-4
- Allison, P. D. (2003). Missing data techniques for structural equation modeling. *The Journal of Abnormal Psychology, 112*, 545-557. doi:10.1037/0021-843X.112.4.545
- Armstrong-Carter, E., Trejo, S., Hill, L. J. B., Crossley, K. L., Mason, D., & Domingue, B. W. (2020). The earliest origins of genetic nurture: The prenatal environment mediates the association between maternal genetics and child development. *Psychological Science, 31*, 781–791. doi:10.1177/0956797620917209
- Bates, T. C., Maher, B. S., Medland, S. E., McAloney, K., Wright, M. J., Hansell, N. K., . . . Gillespie, N. A. (2018). The nature of nurture: Using a virtual-parent design to test

- parenting effects on children's educational attainment in genotyped families. *Twin Research and Human Genetics*, *21*, 73-83. doi:10.1017/thg.2018.11
- Belsky, D. W., Moffitt, T. E., Corcoran, D. L., Domingue, B., Harrington, H., Hogan, S., . . . Caspi, A. (2016). The genetics of success: How single-nucleotide polymorphisms associated with educational attainment relate to life-course development. *Psychological Science*, *27*, 957-972. doi:10.1177/0956797616643070
- Bollen, K. A., & Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological methodology*, *20*, 115-140. doi:10.2307/271084
- Borriello, G. A., Ramos, A. M., Natsuaki, M. N., Reiss, D., Shaw, D. S., Leve, L. D., & Neiderhiser, J. M. (2020). The intergenerational transmission of mathematics achievement in middle childhood: A prospective adoption design. *Developmental Science*, *e12974*. doi:10.1111/desc.12974
- Bouchard, T. J., Jr., & McGue, M. (1981). Familial studies of intelligence: A review. *Science*, *212*, 1055-1059. doi:10.1126/science.7195071
- Briley, D. A., & Tucker-Drob, E. M. (2013). Explaining the increasing heritability of cognitive ability across development: A meta-analysis of longitudinal twin and adoption studies. *Psychological Science*, *24*, 1704-1713. doi:10.1177/0956797613478618
- Capron, C., & Duyme, M. (1989). How much can we raise or lower I.Q.? The joint effects genetic and prenatal factors versus postnatal factors. *Journal of Human Genetics*, *37*, 115-118.
- Cheesman, R., Hunjan, A., Coleman, J. R. I., Ahmadzadeh, Y., Plomin, R., McAdams, T. A., . . . Breen, G. (2020). Comparison of adopted and nonadopted individuals reveals gene-environment interplay for education in the UK Biobank. *Psychological Science*, *31*, 582-591. doi:10.1177/0956797620904450

Cioffi, C. C., Griffin, A. M., Natsuaki, M. N., Shaw, D. S., Reiss, D., Ganiban, J. M., . . .

Leve, L. D. (2021). The role of negative emotionality in the development of child executive function and language abilities from toddlerhood to first grade: An adoption study. *Developmental Psychology, 57*, 347-360. doi:10.1037/dev0000972

Davis, O. S., Haworth, C. M., & Plomin, R. (2009a). Dramatic increase in heritability of cognitive development from early to middle childhood: An 8-year longitudinal study of 8,700 pairs of twins. *Psychological Science, 20*, 1301-1308. doi:10.1111/j.1467-9280.2009.02433.x

Davis, O. S., Haworth, C. M., & Plomin, R. (2009b). Learning abilities and disabilities: generalist genes in early adolescence. *Cognitive Neuropsychiatry, 14*, 312-331. doi:10.1080/13546800902797106

Deary, I. J., Weiss, A., & Batty, G. D. (2010). Intelligence and personality as predictors of illness and death: How researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychological Science in the Public Interest, 11*, 53-79. doi:10.1177/1529100610387081

Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review, 108*, 346-369. doi:10.1037/0033-295x.108.2.346

Domingue, B. W., & Fletcher, J. (2020). Separating measured genetic and environmental effects: Evidence linking parental genotype and adopted child outcomes. *Behavior Genetics, 50*, 301-309. doi:10.1007/s10519-020-10000-4

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., . . . Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*, 1428-1446. doi:10.1037/0012-1649.43.6.1428

- Duyme, M., Dumaret, A. C., & Tomkiewicz, S. (1999). How can we boost IQs of "dull children"? A late adoption study. *Proceedings of the National Academy of Sciences of the United States of America*, *96*, 8790-8794. doi:10.1073/pnas.96.15.8790
- Gagne, J. R., Asherson, P., & Saudino, K. J. (2020). A twin study of inhibitory control at age two and ADHD behavior problems at age three. *Behavior Genetics*, *50*, 289-300. doi:10.1007/s10519-020-09997-5
- Galsworthy, M. J., Dionne, G., Dale, P. S., & Plomin, R. (2000). Sex differences in early verbal and non-verbal cognitive development. *Developmental Science*, *3*, 206-215. doi:10.1111/1467-7687.00114
- Ge, X., Natsuaki, M. N., Martin, D. M., Leve, L. D., Neiderhiser, J. M., Shaw, D. S., . . . Reiss, D. (2008). Bridging the divide: Openness in adoption and postadoption psychosocial adjustment among birth and adoptive parents. *The Journal of Family Psychology*, *22*, 529-540. doi:10.1037/a0012817
- Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3 1/2-7 years old on a Stroop-like day-night test. *Cognition*, *53*, 129-153. doi:10.1016/0010-0277(94)90068-x.
- Goldsmith, H. H., & Rothbart, M. K. (1999). *The laboratory temperament assessment battery (LAB-TAB): Prelocomotor version, technical manual* (3.1 ed.). Madison, WI: Department of Psychology, University of Wisconsin.
- Good, R. H., & Kaminski, R. A. (2002). *Dynamic indicators of basic early literacy skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.
- Haworth, C. M., Wright, M. J., Luciano, M., Martin, N. G., de Geus, E. J., van Beijsterveldt, C. E., . . . Plomin, R. (2010). The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular Psychiatry*, *15*, 1112-1120. doi:10.1038/mp.2009.55

- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55. doi:10.1080/10705519909540118
- Hummer, R. A., & Hernandez, E. M. (2013). The effect of educational attainment on adult mortality in the United States. *Population Bulletin, 68*, 1-16.
- Kaminski, R. A., & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review, 25*, 215-227.
doi:10.1080/02796015.1996.12085812
- Kaufman, A. S., & Lichtenberger, E. O. (1999). *Essentials of WAIS-III assessment*. New York, NY, US: John Wiley & Sons.
- Kendler, K. S., & Neale, M. C. (2010). Endophenotype: A conceptual analysis. *Molecular Psychiatry, 15*, 789-797. doi:10.1038/mp.2010.8
- Kendler, K. S., Turkheimer, E., Ohlsson, H., Sundquist, J., & Sundquist, K. (2015). Family environment and the malleability of cognitive ability: A Swedish national home-reared and adopted-away cosibling control study. *Proceedings of the National Academy of Sciences of the United States of America, 112*, 4612-4617.
doi:10.1073/pnas.1417106112
- Klee, T., Carson, D. K., Gavin, W. J., Hall, L., Kent, A., & Reece, S. (1998). Concurrent and predictive validity of an early language screening program. *Journal of Speech, Language, and Hearing Research, 41*, 627-641. doi:10.1044/jslhr.4103.627
- Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development, 67*, 490-507.

- Kochanska, G., Murray, K. T., & Harlan, E. T. (2000). Effortful control in early childhood: Continuity and change, antecedents, and implications for social development. *Developmental Psychology, 36*, 220-232.
- Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsdottir, B. J., Young, A. I., Thorgeirsson, T. E., . . . Stefansson, K. (2018). The nature of nurture: Effects of parental genotypes. *Science, 359*, 424-428. doi:10.1126/science.aan6877
- Kosik, R., Mandell, G., Fan, A. P., Nguyen, T., Chen, J., & Eaton, W. (2018). The association between childhood educational attainment and adult mental health and status: A thirty-year longitudinal follow up study. *The European Journal of Psychiatry, 32*, 53-62. doi:10.1016/j.ejpsy.2018.01.001
- Kovas, Y., Voronin, I., Kaydalov, A., Malykh, S. B., Dale, P. S., & Plomin, R. (2013). Literacy and numeracy are more heritable than intelligence in primary school. *Psychological Science, 24*, 2048-2056. doi:10.1177/0956797613486982
- Leve, L. D., DeGarmo, D. S., Bridgett, D. J., Neiderhiser, J. M., Shaw, D. S., Harold, G. T., . . . Reiss, D. (2013). Using an adoption design to separate genetic, prenatal, and temperament influences on toddler executive function. *Developmental Psychology, 49*, 1045-1057. doi:10.1037/a0029390
- Leve, L. D., Neiderhiser, J. M., Ganiban, J. M., Natsuaki, M. N., Shaw, D. S., & Reiss, D. (2019). The Early Growth and Development Study: A dual-family adoption study from birth through adolescence. *Twin Research and Human Genetics, 1-12*. doi:10.1017/thg.2019.66
- Leve, L. D., Neiderhiser, J. M., Shaw, D. S., Ganiban, J., Natsuaki, M. N., & Reiss, D. (2013). The Early Growth and Development Study: A prospective adoption study from birth through middle childhood. *Twin Research and Human Genetics, 16*, 412-423. doi:10.1017/thg.2012.126

- Lonigan, C. J., Allan, N. P., & Lerner, M. D. (2011). Assessment of preschool early literacy skills: Linking children's educational needs with empirically supported instructional activities. *Psychology in the Schools, 48*, 488-501. doi:10.1002/pits.20569
- Lonigan, C. J., Wagner, R. K., & Torgesen, J. K. (2007). *Test of preschool early literacy examiner's manual*. Austin, TX: ProED.
- Lugo-Gil, J., & Tamis-LeMonda, C. S. (2008). Family resources and parenting quality: Links to children's cognitive development across the first 3 years. *Child Development, 79*, 1065-1085. doi:10.1111/j.1467-8624.2008.01176.x
- Marceau, K., De Araujo-Greecher, M., Miller, E. S., Massey, S. H., Mayes, L. C., Ganiban, J. M., . . . Neiderhiser, J. M. (2016). The perinatal risk index: Early risks experienced by domestic adoptees in the United States. *PLoS One, 11*, e0150486. doi:10.1371/journal.pone.0150486
- McClelland, M. M., Acock, A. C., Piccinin, A., Rhea, S. A., & Stallings, M. C. (2013). Relations between preschool attention span-persistence and age 25 educational outcomes. *Early Childhood Research Quarterly, 28*, 314-324. doi:10.1016/j.ecresq.2012.07.008
- McNeil, T. F., Cantor-Graae, E., & Sjostrom, K. (1994). Obstetric complications as antecedents of schizophrenia: Empirical effects of using different obstetric complication scales. *Journal of Psychiatric Research, 28*, 519-530. doi:10.1016/0022-3956(94)90042-6
- Montgomery, D. E., & Koeltzow, T. E. (2010). A review of the day–night task: The Stroop paradigm and interference control in young children. *Developmental Review, 30*, 308-330. doi:10.1016/j.dr.2010.07.001

- Neiss, M., & Rowe, D. C. (2000). Parental education and child's verbal IQ in adoptive and biological families in the National Longitudinal Study of Adolescent Health. *Behavior Genetics, 30*, 487-495. doi:10.1023/a:1010254918997
- Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition, 19*, 625-664. doi:10.1521/soco.19.6.625.20886
- Plomin, R. (1986). Multivariate analysis and developmental behavioral genetics: Developmental change as well as continuity. *Behavior Genetics, 16*, 25-43. doi:10.1007/BF01065477
- Plomin, R. (2018). *Blueprint: How DNA makes us who we are*. London: Allen Lane/Penguing Press.
- Plomin, R., Fulker, D., Corley, R., & DeFries, J. (1997). Nature, nurture, and cognitive development from 1 to 16 years: A parent-offspring adoption study. *Psychological Science, 8*, 442-447. doi:10.1111/j.1467-9280.1997.tb00458.x
- Plomin, R., & Kovas, Y. (2005). Generalist genes and learning disabilities. *Psychological Bulletin, 131*, 592-617. doi:10.1037/0033-2909.131.4.592
- Polderman, T. J., Posthuma, D., De Sonneville, L. M., Stins, J. F., Verhulst, F. C., & Boomsma, D. I. (2007). Genetic analyses of the stability of executive functioning during childhood. *Biological Psychology, 76*, 11-20. doi:10.1016/j.biopsycho.2007.05.002
- Rescorla, L., & Alley, A. (2001). Validation of the language development survey (LDS): A parent report tool for identifying language delay in toddlers. *Journal of Speech, Language, and Hearing Research, 44*, 434-445. doi:10.1044/1092-4388(2001/035)
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 42*, 1-36. doi:10.18637/jss.v048.i02

- Saccuzzo, D. P., Johnson, N. E., & Russell, G. (1992). Verbal versus performance IQs for gifted African-American, Caucasian, Filipino, and Hispanic children. *Psychological Assessment, 4*, 239–244. doi:10.1037/1040-3590.4.2.239
- Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype → environment effects. *Child Development, 54*, 424–435.
doi:10.2307/1129703
- Selzam, S., Dale, P. S., Wagner, R. K., DeFries, J. C., Cederlof, M., O'Reilly, P. F., . . . Plomin, R. (2017). Genome-wide polygenic scores predict reading performance throughout the school years. *Scientific Studies of Reading, 21*, 334-349.
doi:10.1080/10888438.2017.1299152
- Sim, F., Thompson, L., Marryat, L., Ramparsad, N., & Wilson, P. (2019). Predictive validity of preschool screening tools for language and behavioural difficulties: A PRISMA systematic review. *PLoS One, 14*, e0211409. doi:10.1371/journal.pone.0211409
- Stromswold, K. (2001). The heritability of language: A review and metaanalysis of twin, adoption, and linkage studies. *Language, 77*, 647-723. doi:10.1353/lan.2001.0247
- Tucker-Drob, E. M., & Bates, T. C. (2016). Large cross-national differences in gene x socioeconomic status interaction on intelligence. *Psychological Science, 27*, 138-149.
doi:10.1177/0956797615612727
- Tucker-Drob, E. M., & Briley, D. A. (2014). Continuity of genetic and environmental influences on cognition across the life span: A meta-analysis of longitudinal twin and adoption studies. *Psychological Bulletin, 140*, 949-979. doi:10.1037/a0035893
- Tucker-Drob, E. M., & Harden, K. P. (2012). Early childhood cognitive development and parental cognitive stimulation: Evidence for reciprocal gene-environment transactions. *Developmental Science, 15*, 250-259. doi:10.1111/j.1467-7687.2011.01121.x

- Wechsler, D. (1997). *WAIS-III administration and scoring manual*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2002). *WPPSI-III administration and scoring manual*. San Antonio, TX: Psychological Corporation.
- Wertz, J., Moffitt, T. E., Agnew-Blais, J., Arseneault, L., Belsky, D. W., Corcoran, D. L., . . . Caspi, A. (2020). Using DNA from mothers and children to study parental investment in children's educational attainment. *Child Development, 91*, 1745-1761.
doi:10.1111/cdev.13329
- Wolchik, S. A., Tein, J. Y., Sandler, I. N., & Kim, H. J. (2016). Developmental cascade models of a parenting-focused program for divorced families on mental health problems and substance use in emerging adulthood. *Development and Psychopathology, 28*, 869-888. doi:10.1017/S0954579416000365
- Woodcock, R., McGrew, K., & Mather, N. (2001). *Woodcock-Johnson tests of achievement (WJ III®)*. Rolling Meadows, IL: Riverside Publishing.
- Yew, S. G., & O'Kearney, R. (2013). Emotional and behavioural outcomes later in childhood and adolescence for children with specific language impairments: Meta-analyses of controlled prospective studies. *Journal of Child Psychology and Psychiatry, 54*, 516-524. doi:10.1111/jcpp.12009
- Yu, H., McCoach, D. B., Gottfried, A. W., & Gottfried, A. E. (2018). Stability of intelligence from infancy through adolescence: An autoregressive latent variable model. *Intelligence, 69*, 8-15. doi:10.1016/j.intell.2018.03.011
- Zilanawala, A., Kelly, Y., & Sacker, A. (2016). Ethnic differences in longitudinal latent verbal profiles in the millennium cohort study. *European Journal of Public Health, 26*, 1011-1016. doi:10.1093/eurpub/ckw184

[Table 1 goes here]

Figure 1*Latent and Observed Variables Used in the Structural Equation Models*

[Figure 1.jpg goes here]

Note. Latent variables are displayed in bold and observed variables are displayed in regular text. BM = birth mother. BF = birth father. WAIS = Wechsler Adult Intelligence Scale-III. WJ = Woodcock-Johnson Tests of Achievement III. LDS = Language Development Scale. AP1 = adoptive parent 1. AP2 = adoptive parent 2. TOPEL = Tests of Preschool Early Literacy. WPPSI = Wechsler Preschool and Primary Scale of Intelligence III. DIBELS = Dynamic Indicators of Basic Early Literacy Skills.

Figure 2

Longitudinal Structural Equation Model Testing the Main Effects of Birth Mother Intellectual Performance on Child Executive Function and Academic Test Performance

[Figure 2.jpg goes here]

Note. Model fit: $\chi^2(170) = 347.59, p < .001$, CFI = .90 RMSEA = .05, SRMR = .06. Standardized estimates reported. Dashed lines represent parameters that are fixed to 1. Adoption openness, child sex, and obstetric risk were included as covariates in the model. BM = birth mother; EF = executive function; WAIS Info = Wechsler Adult Intelligence Scale-III Information Subscale; WJ = Woodcock-Johnson III; LW = letter-word association; RF = reading fluency; WA = word-attack; MF = math fluency; GD = gift delay; FG = forbidden gift; GG = guessing game; DT = dinky toys; G NG = Go NoGo.

*** $p < .001$. ** $p < .01$. * $p < .05$. † $p < .1$. ^{ns} $p \geq .1$.

Figure 3

Longitudinal Structural Equation Model Testing the Main Effects of Birth Father Intellectual Performance on Child Executive Function and Academic Test Performance

[Figure 3.jpg goes here]

Note. Model fit: $\chi^2(170) = 347.59, p < .001$, CFI = .90 RMSEA = .05, SRMR = .06. Standardized estimates reported. Dashed lines represent parameters that are fixed to 1. Adoption openness, child sex, and obstetric risk were included as covariates in the model. BF = birth father; EF = executive function; WAIS Info = Wechsler Adult Intelligence Scale-III Information Subscale; WJ = Woodcock-Johnson III; LW = letter-word association; RF = reading fluency; WA = word-attack; MF = math fluency; GD = gift delay; FG = forbidden gift; GG = guessing game; DT = dinky toys; G NG = Go NoGo.

*** $p < .001$. ** $p < .01$. * $p < .05$. † $p < .1$. ^{ns} $p \geq .1$.

Figure 4

Longitudinal Structural Equation Model Testing Main Effects of Birth Mother Intellectual Performance on Child Verbal Performance and Academic Test Performance

[Figure 4.jpg goes here]

Note. Model fit: $\chi^2(190) = 403.06$, $p < .001$, CFI = .91, RMSEA = .05, SRMR = .07. Standardized estimates reported. Dashed lines represent parameters that are fixed to 1. Adoption openness, child sex, and obstetric risk were included as covariates in the model. BM = birth mother; WAIS Info = Wechsler Adult Intelligence Scale-III Information Subscale; WJ = Woodcock-Johnson III; LW = letter-word association; RF = reading fluency; WA = word-attack; MF = math fluency; LDS = Language Development Scale; AP1 = adoptive parent 1 report; AP2 = adoptive parent 2 report; PK = Test of Preschool Early Literacy (TOPEL) Print Knowledge; DV = TOPEL Definitional Vocabulary; PA = TOPEL Phonological Awareness; WPPSI = Wechsler Preschool and Primary Scale of Intelligence III; ISF = Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Initial Sound Fluency; LNF = DIBELS Letter Naming Fluency; PSF = DIBELS Phoneme Segmentation Fluency; NWF = DIBELS Nonsense Word Fluency.

*** $p < .001$. ** $p < .01$. * $p < .05$. † $p < 0.1$. ^{ns} $p > 0.1$.

Figure 5

Longitudinal Structural Equation Model Testing Main Effects of Birth Father Intellectual Performance on Child Verbal Performance and Academic Test Performance

[Figure 5.jpg goes here]

Note. Model fit: $\chi^2(156) = 309.54$, $p < .001$, CFI = .91, RMSEA = .05, SRMR = .08. Standardized estimates reported. Dashed lines represent parameters that are fixed to 1. Adoption openness, child sex, and obstetric risk were included as covariates in the model. BF = birth father; WAIS Info = Wechsler Adult Intelligence Scale-III Information Subscale; WJ = Woodcock-Johnson III; LW = letter-word association; RF = reading fluency; WA = word-attack; MF = math fluency; PK = Test of Preschool Early Literacy (TOPEL) Print Knowledge; DV = TOPEL Definitional Vocabulary; PA = TOPEL Phonological Awareness; WPPSI = Wechsler Preschool and Primary Scale of Intelligence; ISF = Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Initial Sound Fluency; LNF = DIBELS Letter Naming Fluency; PSF = DIBELS Phoneme Segmentation Fluency; NWF = DIBELS Nonsense Word Fluency.

*** $p < .001$. ** $p < .01$. * $p < .05$. † $p < 0.1$. ^{ns} $p > 0.1$.