

This is a repository copy of *Do Children Cause the Cognitive Stimulation they Receive? Modelling the Direction of Causality*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/229374/>

Version: Published Version

---

**Article:**

Starr, Alexandra, Oginni, Olakunle and von Stumm, Sophie orcid.org/0000-0002-0447-5471 (2024) Do Children Cause the Cognitive Stimulation they Receive? Modelling the Direction of Causality. Behavior genetics. pp. 443-455. ISSN 0001-8244

<https://doi.org/10.1007/s10519-024-10195-w>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



# Do Children Cause the Cognitive Stimulation they Receive? Modelling the Direction of Causality

Alexandra Starr<sup>1</sup> · Olakunle Oginni<sup>2,3</sup> · Sophie von Stumm<sup>1</sup>

Received: 18 December 2023 / Accepted: 6 August 2024 / Published online: 10 September 2024  
© The Author(s) 2024

## Abstract

We tested the directionality of associations between children's early-life cognitive development and the cognitive stimulation that they received from their parents. Our sample included up to 15,314 children from the Twins Early Development Study (TEDS), who were born between 1994 and 1996 in England and Wales and assessed at ages 3 and 4 years on cognitive development and cognitive stimulation, including singing rhymes, reading books, and playing games. Using genetically informative cross-lagged models, we found consistent, bidirectional effects from cognitive development at age 3 to cognitive stimulation at age 4, and from cognitive stimulation at age 3 to cognitive development at age 4. These cross-lagged longitudinal effects were largely explained by underlying common genetic and shared environmental factors, rather than reflecting causal mechanisms. Our findings emphasize the active role that children play in constructing their own learning experiences.

**Keywords** Cognitive development · Cognitive stimulation · Cross-lagged · Genetics · Polygenic scores

## Introduction

Children's early-life cognitive development – the ability to think, reason, and learn – is a powerful predictor of later developmental outcomes like school performance and well-being (Feinstein and Bynner 2004; von Stumm et al. 2020). While the causes for children's differences in cognitive development are manifold, a great focus has been on parents' efforts to engage children in activities that promote learning, known as cognitive stimulation. Positive associations have been widely reported between children's

cognitive development and their engagement in cognitively stimulating activities, such as shared book reading (Xie et al. 2018), conversing with children (Hart and Risley 1995), and singing nursery rhymes together (Mullen 2017). However, the evolution of the association between cognitive development and cognitive stimulation remains unclear in two ways. First, we do not know to what extent cognitive development and stimulation may cause each other, rather than having a common origin, for example when the same genetic or environmental factors affect both constructs (cf. genetic confounding; Wertz et al. 2019; Fig. 1, panel (i)). Many studies have shown that so-called environmental measures are genetically influenced and not truly exogenous to the individual (Dick 2011; Krapohl et al. 2017; Plomin 1995). Furthermore, the environmental and genetic factors that influence one phenotype tend to also influence related phenotypes (Avinun 2020). It follows that individual differences in phenotypically related constructs – like cognitive development and cognitive stimulation – often share common etiologies.

Second, if there is an independent causal relationship between cognitive development and stimulation, what is its direction? At a first glance, we might assume that cognitive stimulation should foster cognitive development, because cognitive stimulation offers opportunities for children to

---

Edited by Tinca Polderman.

✉ Alexandra Starr  
alexandra.starr@york.ac.uk

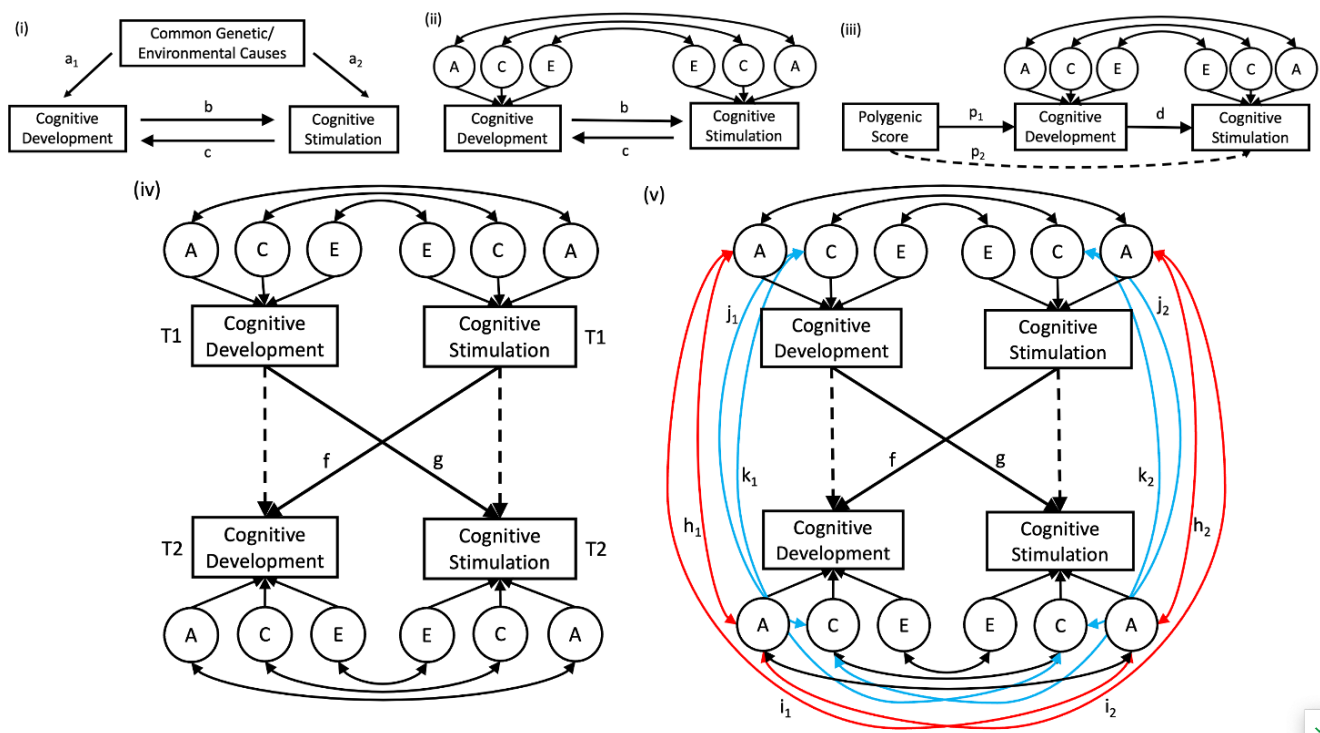
Olakunle Oginni  
OginniO@cardiff.ac.uk

Sophie von Stumm  
sophie.vonstumm@york.ac.uk

<sup>1</sup> University of York, York, UK

<sup>2</sup> King's College London, London, UK

<sup>3</sup> Wolfson Centre for Young People's Mental Health, Cardiff University, Cardiff, UK



**Fig. 1** Genetically informed models for testing the direction of causality between cognitive development and cognitive stimulation

**Note** Panel (i) illustrates that common causes can account for the relationship between cognitive development and cognitive stimulation (paths  $a_1/2$ ), or both constructs may be independently causally associated in either direction (b or c). Panel (ii) maps the Direction-of-Causality (DoC) model, which uses the cross-trait, cross-twin correlations of additive genetic (A), shared environmental (C), and non-shared environmental (E) factors to infer the direction of causality (b or c). The Mendelian Randomization (MR) of the DoC (panel (iii)) uses polygenic scores (PGS) as the instrument variable ( $p_1$ ) to test a

unidirectional causal path, here from cognitive development to cognitive stimulation (d), while controlling for direct pleiotropy ( $p_2$ , dotted arrow). The cross-lagged twin model (panel (iv)) tests causal influences across T1 and T2 (cross-lagged paths f and g) while controlling for the constructs' stability (dotted arrows) and ACE correlations within timepoints. The extended cross-lagged twin model in panel (v) tests for the possibility that the direct, cross-lagged paths in panel (iv) can be attributed to cross-time correlations of genetic factors ( $r_A$  in red colour;  $h_1/2$  and  $i_1/2$ ) and shared environmental factors ( $r_C$  in blue colour;  $j_1/2$  and  $k_1/2$ )

grow cognitively (Tucker-Drob and Harden 2012; Xiong et al. 2020). However, the reverse is also possible: children with better cognitive abilities may elicit or 'cause' greater cognitive stimulation from their environments. For example, precocious children might evoke more learning opportunities from their parents, like being taken to the library, visiting museums, or having frequent and challenging conversations.

Genetically informative study designs can help disentangle relationships between cognitive development and cognitive stimulation (McAdams et al. 2021). The classical twin design compares the phenotypic resemblance of monozygotic (MZ) twin pairs (who share 100% of their genome and are genetically identical) to that of dizygotic (DZ) twin pairs (who, on average, share 50% of their segregating genes) to estimate the respective contributions of genetic and environmental influences to phenotypic trait differences. Twin models can be used to decompose variance in phenotypic traits into additive genetic (A), shared environmental (C), and unique environmental influences

(E; ACE models). An extension of the ACE model is the Direction of Causality (DoC) model that tests for bidirectional causal relationships between two constructs, if certain assumptions are met (Heath et al. 1993; Fig. 1 panel (ii)). Specifically, DoC models can ascertain the direction of causality in cross-sectional data for two traits, whose mode of inheritance is sufficiently different (e.g., A on one trait is significantly greater or lower than on the other). For example, DoC models recently showed that 7-year-olds' ability to read determines how much they choose to read, rather than the degree of print exposure determining their ability to read (van Bergen et al. 2018).

An extension of the DoC model is the Mendelian-Randomization Direction of Causation (MRDoC) model that includes polygenic scores (PGS) as an instrument variable in twin samples for whom genotype data are available (Minică et al. 2018; Fig. 1 panel (iii)). This model tests unidirectional rather than bidirectional causality, while relaxing the assumptions of the DoC models, meaning that the mode of inheritance does not have to differ for both constructs. PGS

aggregate thousands of DNA variants, whose associations with a target phenotype were identified in genome-wide association (GWA) studies, and index individuals' genetic propensities for phenotypic development (e.g., Plomin & von Stumm, 2018). PGS can serve as instrument variables in MR models if they are plausible and significant predictors of the exposure, whose causal effect on an outcome is being tested (Sanderson et al. 2022; Davey Smith and Ebrahim 2003). Because no GWA study on children's early life cognitive development has been published to date, studies that seek to capture children's genetic propensities for cognitive development utilise PGS based on a GWA study for years spent in education (e.g., von Stumm et al. 2023). Educational attainment and cognitive traits are genetically correlated phenotypes (e.g., Calvin et al. 2012), even though the extent to which the DNA variants associated with years spent in education are identical with those for early life cognitive development is unclear. Thus, PGS for years spent in education can be used in an MRDoC model to test if cognitive development causes cognitive stimulation (Fig. 1, panel (iii)). Because no GWA study has been reported for cognitive stimulation, testing the reverse direction of causality from cognitive stimulation to cognitive development is not possible (i.e., there is no PGS for cognitive stimulation).

A third option is the cross-lagged twin model (Burt et al. 2005), which tests the effect of a trait at an earlier assessment time (T1) on another trait at a later assessment time (T2; i.e., cross-lagged paths), while controlling for the constructs' stability over time and their ACE factors' correlations within the assessment times (Fig. 1, panel (iv)). Cross-lagged twin models do not require the constructs to have different modes of inheritance, and they do not include instrument variables. Yet, they necessitate longitudinal data, with the two constructs of interest being assessed two or more times. The cross-lagged twin model can be further extended to control for the cross-time correlations of the genetic and shared environmental influences on cognitive development and cognitive stimulation. This extended cross-lagged twin model allows testing if the direct effects across time – that is, the longitudinal stability of each construct and cross-lagged effects between the constructs – are confounded by genetic and shared environmental factors that affect both constructs. In other words, this model (Fig. 1, panel (v)) tests if genetic or shared environmental factors act as a common origin of cognitive development and cognitive stimulation regarding their cross-sectional as well as longitudinal relationship. Accounting for these potential confounding effects strengthens the rationale for inferring causal effects from the direct paths in the model compared to alternative approaches (Burt et al. 2005).

The aim of the current research was to clarify the direction of causality of the association between children's

cognitive development and the cognitive stimulation that they receive. We fitted a series of genetically informative models in a large UK-population representative twin study with genotype data and measures of cognitive development and cognitive stimulation, including singing rhymes, reading books, and playing games, at the ages of 3 and 4 years.

## Methods

### Sample

The sample was drawn from the Twins Early Development Study (TEDS) that recruited 16,810 families with twin children born between 1994 and 1996 in England and Wales (Rimfeld et al. 2019). TEDS families were representative of other UK families in the 1990s; for example, 93% of the TEDS families identified as white vs. 93% of the UK's families, 38% of TEDS mothers (47% of fathers) attained A-levels or a higher educational qualification vs. 35% of mothers in the UK population (47% of fathers), and 44% of TEDS mothers (92% of fathers) reported being employed at twins' ages 2–4 years vs. 50% of UK mothers (91% fathers; Rimfeld et al. 2019). At the twins' ages 3 and 4 years, 9,350 and 12,528 families were contacted respectively for data collection, of whom 6,119 and 8,198 families responded, providing data for 12,118 and 16,303 individual children at both timepoints (see the TEDS website for further details: <https://www.teds.ac.uk/datadictionary/studies/returns/samples.htm>). The TEDS project approval (05.Q0706/228) was granted by the ethics committee for the Institute of Psychiatry, Psychology and Neuroscience at King's College London.

We excluded 1,797 twins because they reported severe medical problems and adverse perinatal conditions, or missed information on zygosity or gender or learning disability, or had a confirmed learning disability (see Table S1 in Supplementary Material). Sample sizes with phenotypic data ranged from a minimum of 10,634 at age 3 to a maximum of 15,314 participants at age 4 years; genotype data was available for 5,185 individuals (see Table S2 for a sample-size-by-variable breakdown and twin pair numbers).

## Measures

### Cognitive Development

Parent-administered tests and parent-reported observations were used to assess twins' cognitive ability at the ages 3 and 4 years. These measures have been validated against standard tests administered by trained testers (Oliver et al. 2002; Saudino et al., 1998). Specifically, nonverbal cognitive

performance was assessed using age-appropriate versions of the Parent Report of Children's Abilities (PARCA; Oliver et al. 2002; Saudino et al., 1998), while verbal ability measures were assessed by parent reports of children's vocabulary and grammar using the CDI–III, an extension of the short form of the MacArthur Communicative Development Inventories: Words and Sentences (Fenson et al., 2007). The PARCA is an established, valid, and reliable measure of children's early-life cognitive abilities at the ages 2, 3, and 4 years (Bayley 1993; Blaggan et al. 2014; d'Apice et al., 2019; Martin et al. 2013; McCarthy 1972; Oliver et al. 2002; Saudino et al., 1998; Price 2002; Oliver and Plomin 2007). The UK's National Institute for Clinical Excellence (NICE 2017) uses a revised version (PARCA-R) in their developmental assessment guidelines. At age 2 years, the verbal subtest of the PARCA comprised one item on grammar (0=talking in short, incomplete sentences to 2=talking in long sentences and using '-est' words and 'but'). The non-verbal subtest comprised five parent-administered items, including twin block building, copying (i.e., follow the leader), drawing, paper folding, and matching. At ages 3 and 4 years, non-verbal ability was assessed with three parent-administered tests, specifically Odd One Out (16 items; e.g., child is asked to point at the two out of three images that go together), drawing (6 items; e.g., parent draws a vertical line and the child is asked to copy it), and puzzle (12 items; e.g., child is asked to identify the next image in a sequence). The parent-administered PARCA component was supplemented by parent-report items on concrete behaviours (e.g., "Does your child ever play any game with another child that involves taking turns?"; 1=yes, 0=no or don't know). At each age, test scores were standardized and summed to a total score.

### Cognitive Stimulation

At twins' ages 3 and 4 years, parents answered the questions from the Home Observation for Measurement of the Environment (HOME) measure (Caldwell and Bradley 1984) for their first-born twin and then rated the degree to which their answer was also true for the second born twin. Answers were given on 5-point Likert scales that referred to frequency of occurrence (1=less than once a month to 5=almost daily) except where indicated otherwise. The HOME Inventory was designed to measure the quantity and quality of stimulation, support, and structure available in children's home environment (Bradley 2015). Out of a total of 17 items asking parents to describe 'twins' early experiences', we identified seven items that mapped onto cognitive stimulation. These seven items are listed below and capture 'involvement in enriching activities'. They were presented to the parents under the headers "How you

talk to your twins" and "Your twins' play". Thus, parents were asked to answer the questions from their perspective of engagement with the children.

1. Does your 1st born twin take part in nursery rhymes, simple songs, or prayers?
2. Does your 1st born twin read books or look at books with you?
3. Do you talk to your 1st born twin when you are doing household chores?
4. Does your 1st born twin have any puzzles (for example, jigsaws, puzzle boards)? Scored from none through 11 or more (5-point scale).
5. Does your 1st born twin have any children's tapes/records/CDs (for example, of nursery rhymes, stories)? Scored from none through 11 or more (5-point scale).
6. Does your 1st born twin have any children's books? Scored from none through 101 or more (5-point scale).
7. Does your 1st born twin play board games or card games with you (for example, "snakes and ladders", "happy families", "snap")?

These items were grouped by themes into Talking and Rhymes (TR; items 1, 3, 5); Playing with Books (PB; items 2 and 6) and Playing Games (PG; items 4 and 7), which may relate differently to cognitive development with differing shares of common genetic and environmental contributions (e.g., Noble et al. 2019; Dowdall et al. 2020). Scores for these domains were derived by summing the respective items at each assessment age.

### Polygenic Scores (PGS) for Years Spent in Education

Saliva and buccal cheek swab samples were collected when the twins were aged either 12 or 16 year. DNA was extracted and genotyped to compute polygenic scores (PGS) based on the latest genome-wide association (GWA) study for years spent in education (Lee et al. 2018; see Supplementary Methods S1 and S2 for details). The PGS were adjusted for the first ten principal components, chip, batch, and plate effects using the regression method.

### Statistical Analyses

Data preparation was carried out using STATA, and OpenMx in R (Neale et al. 2016) was used for subsequent analyses, which were preregistered at <https://osf.io/5ybg4/>. We departed in several ways from the preregistration. A first change was that we excluded the cognitive ability at age 2 years from our analyses, which would have violated the temporality in exploring bidirectional causal relationships.



We specified one latent factor for cognitive development (CD) with the cognitive ability test scores at ages 3 and 4 years as indicators to minimize the impacts of measurement error in preliminary genetic analyses and in the Direction of Causation biometric genetic model which can be biased by measurement error (Heath et al. 1993). Specifically, three latent factors for the cognitive stimulation domains including Talk and Rhyme ( $CS_{TR}$ ), Play with books ( $CS_{PB}$ ) and Play Games ( $CS_{PG}$ ) each indicated by their respective subdomain scores at ages 3 and 4 years respectively; and a latent Genetic Propensity for Education Attainment ( $GP_{EA}$ ) factor with the PGS for years spent in education (Lee et al. 2018) as the single indicator. To scale the latent factors, one factor loading per factor was fixed to 1. The residual variances of the factor with two indicators were constrained to be equal to identify the measurement models; the residual variance for the single-indicator factors was fixed at 0. All the observed indicator variables were residualized for age and sex. For the other genetic models (i.e., Mendelian Randomization-Direction of Causation and Cross-lagged models), we used the observed variables.

### Phenotypic Analyses

Correlations between the latent factors of  $GP_{EA}$ , CD,  $CS_{TR}$ ,  $CS_{PB}$  and  $CS_{PG}$  were estimated using Full Information Maximum Likelihood (FIML) estimation that utilises all the available data points for analyses (Enders and Bandalos 2001), including within-individual correlations that were constrained to be equal across birth order and zygosity, and cross-twin correlations which were estimated separately for monozygotic and dizygotic twins. Based on previous studies (Polderman et al. 2015), we expected the monozygotic cross-twin correlations to be greater than, but less than twice, the dizygotic cross-twin correlations to be consistent with models including additive genetic (A), shared/common environmental (C) and non-shared environmental (E) influences.

### Genetic Models

The classical twin design parses the variances and covariances of variables into A, C and E components by comparing correlations between monozygotic and dizygotic twin pairs. This rests on the assumption that twin pairs who grew up together are influenced by their shared environments to the same extents, that genetic and environmental influences are distinct and that mate selection is random (Neale and Cardon 1992; Rijdsdijk and Sham 2002). Univariate models were specified to estimate A, C and E influences on each of the latent phenotypes (i.e., CD,  $CS_{TR}$ ,  $CS_{PB}$  and  $CS_{PG}$ ). The

variance in  $GP_{EA}$  was not similarly parsed because all its variance is assumed to reflect genetic influences.

**Direction of Causation (DoC) Model** Three separate bivariate DoC models were specified between CD and each of the three cognitive stimulation factors  $CS_{TR}$ ,  $CS_{PB}$  and  $CS_{PG}$ . The full DoC model specifies ACE variance components for each latent factor and two bidirectional causal paths between the two factors (i.e.,  $CD \rightleftharpoons CS_{TR}/CS_{PB}/CS_{PG}$ ; Heath et al. 1993; Tick et al. 2016). Two further models were fitted, specifying single causal paths (i.e.,  $CD \rightarrow CS_{TR}/CS_{PB}/CS_{PG}$  and  $CS_{TR}/CS_{PB}/CS_{PG} \rightarrow CD$  respectively). Each full, bidirectional DoC model, and the unidirectional DoC models were compared against a bivariate Cholesky decomposition, which partitions all variances and covariances into ACE components. Model comparisons were based on Chi-squared difference tests and the Akaike Information Criterion (AIC), with lower values indicating better model fit (Gillespie et al. 2003; Rasmussen et al. 2019). DoC models require ACE influences on the phenotypes to differ sufficiently (Heath et al. 1993; Rasmussen et al. 2019) but this was not the case in the current analyses. We therefore proceeded to fitting additional genetic cross-lagged models, which were not preregistered (described below).

### Mendelian Randomization-Direction of Causation (MRDoC) Model

The MRDoC incorporates Mendelian Randomization (MR) and DoC approaches to test unidirectional causation using cross-sectional data (Minică et al. 2018). We fitted MRDoC models to test whether cognitive development causally influenced the cognitive stimulation factors at age 3 and 4 years; we could not test whether cognitive stimulation causally influenced cognitive development, because no adequate instrument variable was available. We tested the significance of the correlation between PGS based on the Lee et al. (2018) GWA study for years spent in education and cognitive development, as well as that of their respective associations within the MRDoC, to infer the suitability of PGS as an instrument variable. We encountered two difficulties with the MRDoC models. First, E influences on the cognitive development and cognitive stimulation latent factors were close to zero, which reduces the power of the MRDoC model to determine causality (Kohler et al. 2011). An alternative approach is specifying MRDoC models for observed scores of cognitive development and stimulation rather than for latent factors, with the limitation that the observed scores will include measurement error. When we used observed scores, the association between PGS and cognitive development was not significant, suggesting that PGS were not a valid instrument variable for our analyses.

Thus, we concluded that the cross-lagged twin model was better suited to the current data.

**Cross-Lagged Twin Model** We tested causal associations between observed scores of CD and  $CS_{TR}/CS_{PB}/CS_{PG}$  at the two timepoints of 3 and 4 years. The observed variables were used for these analyses to enable us take advantage of the longitudinal nature of the data. The genetic cross-lagged model (Burt et al. 2005) adjusts for the stability of constructs across assessment times (i.e., autoregressive paths) and for correlations between the ACE influences on each variable within each assessment age. Our extended cross-lagged twin model further controls for cross-time correlations of genetic ( $r_A$ ) and shared environmental ( $r_C$ ) factors that affect cognitive development and cognitive stimulation and might therefore confound potentially causal effects. The unique environmental cross-time correlations were restricted to zero as is commonly done when investigating causal mechanisms in twin modelling (e.g., Minică et al. 2018; Kohler et al. 2011) to facilitate statistical identification which was checked using the *mxCheckIdentification* function in OpenMx (Neale et al. 2016). The extended cross-lagged model we apply here is comparable to the approach introduced by Starr and colleagues (2023) testing cross-time causal effects between household chaos and school performance while controlling for confounding by cross-time genetic and shared environmental correlations. Our analyses applying the cross-lagged twin model and its extension were not preregistered.

## Results

Means, Standard Deviations, and Skewness of the cognitive development (CD) and cognitive stimulation variables (Talking and Rhymes, Playing with Books and Playing Games) are summarized in Table 1. The cross-twin correlations for each of the latent factors among DZ twins were lower but greater than half those in MZs suggesting A and C influences, while small E influences were indicated by the MZ cross-twin correlations being close to 1 (Table 2). The correlations between cognitive development and cognitive stimulation factors were positive and significant, ranging from moderate ( $r=0.38$ , 95% CI: 0.39–0.45; for  $CS_{PB}$ ) to large ( $r=0.51$ , 95% CI: 0.49–0.54; for  $CS_{PG}$ ). Similarly, the PGS, based on Lee et al.'s (2018) GWA study on years spent in full-time education, were significantly correlated with factors of cognitive development ( $r=0.07$ , 95% CI: 0.03–0.10), and with cognitive stimulation ( $r=0.09$ , 0.18, and 0.12 for  $CS_{TR}$ ,  $CS_{PB}$ , and  $CS_{PG}$ , 95% CIs: 0.05–0.22),

**Table 1** Descriptives for cognitive development and cognitive stimulation variables at age 3 and 4 years

	Cognitive development CD						Cognitive stimulation CS					
	PARCA scores			Talking-rhyme $CS_{TR}$			Playing books $CS_{PB}$			Playing games $CS_{PG}$		
	N	Mean	SD	Skew	N	Mean	SD	Skew	N	Mean	SD	Skew
Age 3	11,140	0.00	1	-0.32	10,643	0.02	1.92	-0.71	10,663	0.04	1.54	-0.21
Age 4	15,314	0.00	1	-0.23	14,015	0.03	1.94	-0.68	14,091	0.03	1.56	-0.10
Note PARCA – Parent Report of Children's Abilities; our analyses are based on standardized scores derived and provided by the TEDS team												

**Table 2** Univariate and bivariate factor correlations and univariate and standardized A, C, and E influences on latent factors of cognitive development and cognitive stimulation with 95% confidence intervals

	CD	CS <sub>TR</sub>	CS <sub>PB</sub>	CS <sub>PG</sub>
Univariate cross-twin correlations				
MZ twin pairs	0.99 (0.97, 1.00)	0.94 (0.92, 0.96)	0.96 (0.95, 0.98)	0.95 (0.93, 0.96)
DZ twin pairs	0.86 (0.85, 0.88)	0.79 (0.76, 0.81)	0.76 (0.74, 0.78)	0.78 (0.76, 0.80)
Bivariate correlations with CD				
Within-person	-	0.42 (0.39, 0.45)	0.38 (0.35, 0.41)	0.51 (0.49, 0.54)
Cross-twin				
MZ twin pairs	-	0.42 (0.39, 0.45)	0.37 (0.35, 0.40)	0.49 (0.47, 0.52)
DZ twin pairs	-	0.38 (0.35, 0.41)	0.34 (0.31, 0.37)	0.44 (0.41, 0.47)
Standardized univariate variance components				
$h^2$	0.24 (0.20, 0.29)	0.30 (0.24, 0.37)	0.40 (0.35, 0.46)	0.33 (0.28, 0.39)
$c^2$	0.74 (0.70, 0.78)	0.63 (0.58, 0.69)	0.56 (0.51, 0.61)	0.61 (0.56, 0.66)
$e^2$	0.01 (0.00, 0.03)	0.06 (0.04, 0.08)	0.04 (0.02, 0.05)	0.05 (0.04, 0.07)

Note CD = Cognitive development latent factor, indicated by PARCA scores at ages 3 and 4 years; CS<sub>TR</sub> = Cognitive Stimulation Talking and Rhyming factor, indicated by Talking and Rhyming scores at ages 3 and 4 years; CS<sub>PB</sub> = Cognitive Stimulation Playing with Books factor, indicated by Playing with Books scores at ages 3 and 4 years; CS<sub>PG</sub> = Cognitive Stimulation Playing Games factor, indicated by Playing Games scores at ages 3 and 4 years; MZ = monozygotic; DZ = Dizygotic;  $h^2$ ,  $c^2$  and  $e^2$  = Standardized additive genetic (A), shared (C) and non-shared (E) environmental influences

indicating significant pleiotropy of the instrument (Table S4a-c). PGS also correlated significantly with observed scores of cognitive stimulation at ages 3 and 4 years ( $r = 0.04$  to  $0.13$ , 95% CIs:  $0.01$ – $0.17$ , Table S5) and with cognitive development at age 4 ( $r = 0.06$ , 95% CI:  $0.03$ – $0.09$ ) but not at age 3 ( $r = 0.03$ , 95% CI:  $0.00$ – $0.06$ ; Table S5).

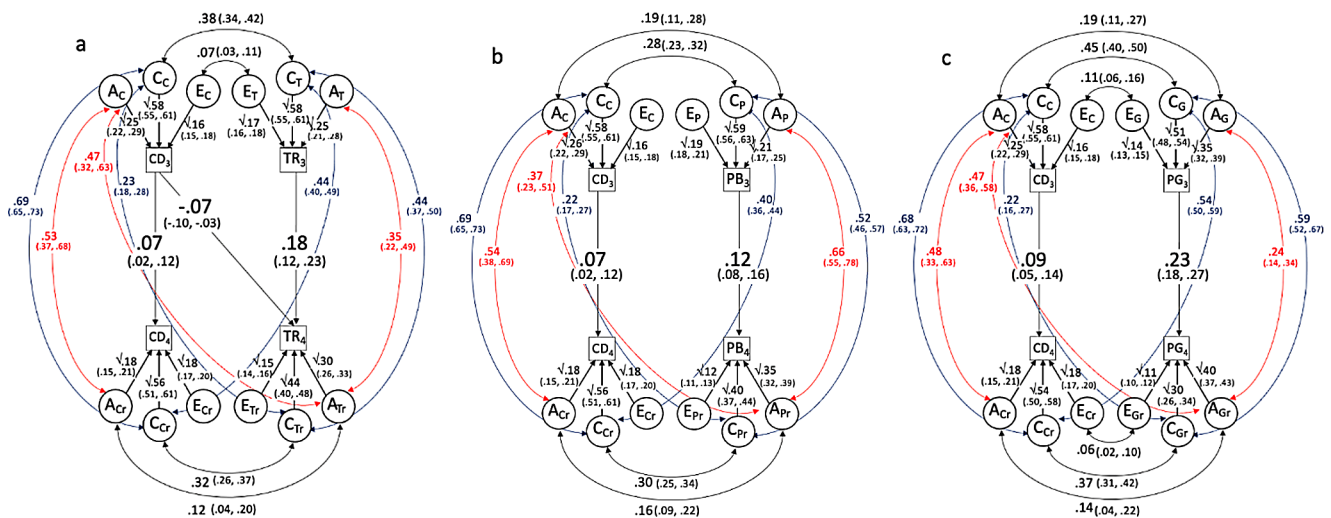
The latent factors (CD, CS<sub>TR</sub>, CS<sub>PB</sub> and CS<sub>PG</sub>) were significantly influenced by latent shared environmental and additive genetic factors (56–74% and 24–40%, respectively), while non-shared environmental influences were minimal (1–6%; Table 2). The covariances between CD, and CS<sub>TR</sub>, CS<sub>PB</sub> and CS<sub>PG</sub> were predominantly attributable to shared environmental influences (ranging from 76 to 83%, Table S3), with relatively smaller contributions from latent additive genetic and non-shared environmental influences (16–20% and 1–4%, respectively). Similarly, the observed variables were mostly influenced by C components (52–63%) while non-shared influences had the smallest effects (9–20%; Table S7).

Based on the similar modes of inheritance for cognitive development and cognitive stimulation, we concluded that the DoC model was not suitable for our data to test causal effects. Likewise, the requirements for the MRDoC model were not met in our data as explained above, thus in the following we focus on the results of the cross-lagged twin models which reflect the best suited modelling approach in our case. For reasons of transparency, we report the results of the DoC models in the Supplemental Materials (see Supplementary Methods S3 for details, Tables S4a-c for model comparisons and Figure S1 for path coefficients). However, we advise against interpreting these because the similar ACE influences – which are required to be significantly different for the Direction of Causation model) – on the cognitive development and cognitive stimulation latent factors may lead to low power to determine the direction of causal effects (Health et al., 1993).

The results from the cross-lagged twin models indicate small, yet significant direct effects of all three CS domains at age 3 on CD at 4 years ( $f = 0.06$ – $0.12$ ) as well as in the reverse direction, i.e., CD at age 3 on CS at age 4 years ( $g = 0.10$ – $0.12$ ), which suggests a bidirectional, causal relationship between cognitive development and stimulation (all model results are included in Figure S2a-c in the Supplemental Materials). However, these results are potentially confounded by cross-time genetic and shared environmental factors common to CD and CS, as these were not accounted for in the model.

We subsequently applied the extended cross-lagged twin model and found that the prospective, bidirectional associations between CD and TR, PB and PG at ages 3 and 4 years were largely due to genetic and shared environmental factors that affected both cognitive development and stimulation across ages. Figure 2a-c displays the best-fitting parsimonious model with all non-significant paths restricted to zero (the full model results are displayed in Figure S3a-c in the Supplemental Materials). The results from the extended cross-lagged model including cross-time genetic and shared environmental correlations also fit the data slightly better compared to the original model by Burt and colleagues (2005; model fit indices and model comparisons reported in Note below Figure S3). After controlling for cross-time  $r_A$  and  $r_C$ , the residual direct cross-lagged effects between CD and TR, PB and PG across age 3 and 4 years were negligible (Fig. 2). An exception was a significant, negative effect of CD at age 3 on TR at age 4 ( $g = -0.07$ ; 95% CI:  $-0.10$ ,  $-0.03$ ). That is, for children with higher cognitive development at age 3, parents reported less engagement in talking and rhyming activities one year later. This negative effect only emerged once common genetic and shared environmental factors explaining the remainder of the positive





**Fig. 2** a–c: Parsimonious genetic cross-lagged models corrected for cross-time genetic and shared environmental correlations

**Note** Models were fitted to observed scores, not latent factors. CD: Cognitive development; TR: Talking and rhyming; PB: Playing with books; PG: Playing games; at ages 3 and 4 years (subscripts 3 and 4, respectively). A, C and E: Additive genetic, shared and unique environmental influences respectively; subscripts C, T, P and G refer to CD, TR, PB and PG respectively, subscript r indicates residual effects at age 4 years. Paths which did not reach significance in the full model

phenotypic association between children's CD at age 3 and the TR stimulation they experience at age 4 were accounted for.

The cross-lagged effects of CD at age 3 on PB and PG at age 4 were largely due to cross-time  $r_A$  and  $r_C$  that explained 28.9% and 47.9%, respectively, of the model induced phenotypic correlation between the two constructs (for all relative contributions of the direct paths and the A, C, and E influences, see Tables S5 and S6). In the reverse direction (i.e.,  $TR_3/PB_3/PG_3 \rightarrow CD_4$ ), the cross-time  $r_C$  accounted for 90.5–94.2% of the overall phenotypic association, while cross-time  $r_A$  explained negligible variance. Regarding both directions (i.e.,  $CD_3 \rightarrow CS_4$  and  $CS_3 \rightarrow CD_4$ ), the remainder of the overall phenotypic association was explained by paths through the within-time A, C, and E correlations at age 3 and the direct stability paths of either construct (0.3–17.9%).

## Discussion

Twin studies can be used to strengthen causal inferences for observed associations between two constructs by controlling for their shared genetic and environmental influences (McAdams et al. 2021). Specifically, twin studies can address whether associations between a putatively environmental exposure and developmental differences in a phenotype remain significant after accounting for the confounding

were restricted to zero and are not shown in the Figure to sustain clarity. Models with non-significant paths restricted to zero did not fit the data worse than the corresponding full models (Supplementary Figure S3a–c;  $X^2_{[4]} = 6.95$ ,  $p = 0.14$ ;  $X^2_{[5]} = 8.75$ ,  $p = 0.12$ ;  $X^2_{[3]} = 4.94$ ,  $p = 0.18$  for models a, b and c respectively). Results from the cross-lagged models not adjusting for cross-time genetic and shared environmental correlations (following the approach by Burt et al. 2005) are presented for comparison in Figure S2a–c in the Supplemental Materials

effects of shared etiology (i.e., common causes; McAdams et al. 2021). Recent years have seen an explosion of novel modelling approaches that extend the classical twin design, which enable causal inferences but are yet to be systematically applied in psychological research (Erbeli et al. 2020; McAdams et al. 2021). Here, we fitted Direction-of-Causation (DoC) models (Heath, 1993), Mendelian Randomization (MR) extensions of the DoC model (Minică et al. 2018), and cross-lagged twin models (Burt et al. 2005; Starr et al. 2023) to investigate if children's cognitive stimulation causes their cognitive development or vice versa at age 3 and 4 years.

We found bidirectional phenotypic associations between children's cognitive development and their cognitive stimulation, which were largely explained by common genetic and shared environmental factors, rather than reflecting causal effects. We only identified one significant direct effect that might be interpreted as a causal path. Cognitive development at age 3 years had a negative effect on the level of engagement in talking and rhyming that parents reported to offer to their children at age 4 years. This – somewhat counterintuitive – finding implies that children with more advanced cognitive development at age 3 years might receive less stimulation from talking and rhyming activities one year later. However, we caution here that this small negative effect only emerged for one of the three cognitive stimulation domains and might be spurious (i.e., false positive finding due to fitting multiple models). We can speculate

about two possible interpretations. First, highly intelligent children may receive less cognitive stimulation over time, perhaps because parents perceive their cognitive stimulation needs as saturated or because parents provide different, more advanced stimulation but reduce rudimentary ones. Second, children who are perceived to show poor cognitive development may receive increased stimulation from their parents to help them advance and catch up. Our models suggest that this small, negative effect of cognitive development on stimulation is likely due to unique environmental factors that are thought to capture unsystematic and idiosyncratic events that make two twins in a family different from one another (Turkheimer and Waldron 2000; Plomin 2011). The exact events, experiences, or factors that drive non-shared environmental variance are yet to be determined, including environmental risk factors, for example families' limited access to resources (i.e., low SES), substantial household chaos, and high levels of parental psychopathology (Asbury et al. 2003). Yet, our findings will need to be replicated and explored further before any conclusions about direct effects between cognitive stimulation and cognitive development can be drawn. Overall, we conclude that simultaneous, bidirectional mechanisms drive the association between cognitive development and stimulation: Genetic factors and environmental experiences shared by two children in a family explain their cognitive development as well as the cognitive stimulation that parents offer and that children evoke.

Our findings indicated that cognitively stimulating experiences are unlikely to exert causal influence on prospective cognitive development or vice versa, after controlling for the common causes of children's cognitive development and cognitive stimulation shared by two children in a family during early childhood (i.e., genetic and shared environmental factors). Randomized controlled trials on the effectiveness of cognitive stimulation interventions to improve children's cognitive development produced mixed results. Some studies found that play and shared book reading interventions might be suitable to improve early life cognitive development (Tachibana et al. 2012; Howard et al., 2017; Dowdall et al. 2020). For example, a meta-analysis of  $n = 54$  studies on the effect of shared book reading on children's language skills showed that effects immediately after the intervention were overall significant with modest in size ( $\bar{g} = 0.194$ ); however, effects tended to be negligible when interventions were tested against active control groups or at follow-up assessments (Noble et al. 2019). Meta-analytic evidence also contends that effects from interventions that successfully raised children's IQ in the short-term tend to fade out over time by about 0.10–0.13 SD per year or were not evident anymore at follow-up assessments conducted less than two years after the end of treatment (Protzko 2015; Takacs and Kassai 2019; Bailey et al. 2020).

Instead, the phenotypic association between earlier cognitive stimulation and later cognitive development is mainly driven by shared environments that likely include parenting behaviors, the family's socioeconomic background, and characteristics of the home environment and wider neighborhood. These environmental factors are shared when they affect both children in the same way (Rijsdijk and Sham 2002). In addition, genetic factors contributed to correlations of cognitive development and stimulation within and across assessment ages. These genetic influences are likely exerted through partly genetically influenced traits, including parents' and children's individual characteristics, such as parental cognitive ability, educational level, and assets as well as children's temperament, their physical health, and complex parent-child interaction processes (Tucker-Drob and Harden 2012; Brownell et al. 2016; Demange et al. 2022). However, the effects of these factors on the association between children's cognitive development and stimulation need to be studied in depth in future research. Apart from potentially driving genetic effects on children's cognitive development, these factors also give rise to passive and evocative gene-environment correlations (Plomin et al. 1977; Price and Jaffee 2008), which may induce bias in the variance estimates in twin models: In the presence of gene-environment correlations genetic effects are underestimated, while the shared environmental component is slightly overestimated (Verhulst and Hatemi 2013). This potential bias may in part explain the strong shared environmental effects (52–63%) and comparatively weaker genetic effects (23–40%) on children's differences in cognitive development and cognitive stimulation in the present study. A similar pattern emerged for the etiologies of the associations between cognitive development and stimulation, which were again mainly due to shared environmental (76–84%) and additive genetic influences (16–20%). Non-shared environmental influences on cognitive development and cognitive stimulation were small and overall negligible. While this study is to our knowledge the first to estimate the heritability of cognitive stimulation and its association with cognitive development, our findings confirm that all traits are heritable, including putatively environmental measures that are genetically influenced (Krapohl et al. 2017; Plomin 1995; Polderman et al. 2015).

Our findings suggest that children's cognitive development is not causally affected by the cognitive stimulation they experience after adjusting for genetic and environmental influences that are shared by two twins within their family. This does not imply, however, that modifying children's cognitive stimulation or other aspects of their home learning environment cannot affect their cognitive development. Behavioural genetic studies can only describe *what is*, not *what could be*; thus, our findings do not rule out the

potential of impact by interventions that target children's cognitive stimulation.

## Limitations

Notwithstanding this study's many strengths, including the analysis of twin and genomic data, large sample sizes, and repeated assessments of the core constructs using state-of-the-art methods, it is not without limitations. First, cognitive stimulation was assessed using seven parent-reported items, but a multi-informant approach (i.e., naturalistic home observations) with a greater number of observed variables would have improved the measures' validity. The cognitive stimulation items were grouped into three themes (i.e., talking and rhymes, playing with books, and playing games), but alternative groupings are possible, such as differentiating cognitive stimulation activities that require initiation by the parents vs. opportunities for stimulation (i.e., toys) that twins can access without direct parent involvement. Second, data on cognitive stimulation were only collected at the twins' ages 3 and 4 years, which made it impossible to test for meaningful changes in association with cognitive development over the longer course of childhood.

Third, our cross-lagged twin models were fitted using observed rather than latent variables. Observed variables are subject to measurement error which can obscure causal effects between cognitive stimulation and development when they are in fact present in the population. In our initial analysis strategy, we specified latent factors for the three domains of cognitive stimulation, for two of which (i.e.,  $CS_{PB}$  and  $CS_{PG}$ ) only two observed indicators were available: for the models to be identified, their factor loadings had to be restricted. As a result, the domains of cognitive stimulation in the cross-lagged twin models included measurement error. That said, the measurement error of cognitive stimulation is likely low here because only very small proportions of nonshared environmental variance, which also capture measurement error, were evident in the univariate decompositions ( $e^2 = 4\text{--}6\%$ ).

Fourth, the two core constructs might have not been sufficiently distinct from each other due to common methods biases, as cognitive stimulation and cognitive development were both reported by parents. Common methods biases can inflate the similarity between twins irrespective of their zygosity (e.g., Neale and Cardon 1992). As a result, the estimate of the shared environment may have been biased upwards for both cognitive development and cognitive stimulation, which could explain the larger shared environmental influences on cognitive development that we observed here than were reported elsewhere (e.g., Grotzinger et al. 2019; Briley and Tucker-Drob 2013), as well as the high level of shared environmental

covariance between constructs. Common methods biases may have also contributed to the genetic architecture of cognitive development and cognitive stimulation not being sufficiently differentiated to meet the assumptions of DoC models (Heath, 1993; van Bergen et al. 2018).

Fifth, the association between the PGS for years spent in education (Lee et al. 2018) and cognitive development was small, which indicates that the PGS was a weak instrument for our phenotypic measure of cognitive development that can lead to an overestimation of causal effects (Burgess and Thompson 2011). The weak associations between the PGS and the phenotype might in this case be due to developmental effects, since the PGS was aimed to capture educational achievement in adulthood. Cognitive development in early childhood differs as a construct from adult educational achievement, is less strongly influenced by genetic factors, and exhibits considerable instability over time compared to cognitive ability in adulthood (von Stumm et al. 2023; Breit et al. 2024; Petrill et al. 2004). As a result, the MRDoC analysis strategy we had preregistered was not viable, and we fitted cross-lagged twin models instead that require longitudinal data.

In sum, maximising the explanatory power of complex genetically informed statistical models, like the ones we attempted to fit here, is often difficult. Leveraging the theoretical value of DoC and MRDoC models is only possible if empirical data are available that meet the essential model assumptions, including sufficiently different aetiologies and valid instrument variables. Finding better synergies between statistical models and available data than reflected in the current study should be a priority for future research.

## Conclusions

Across different genetically sensitive study designs, we found that genetic and shared environmental factors drive the association between children's differences in cognitive development and in the cognitive stimulation that they receive from their parents during the early years. While our study did not produce evidence for causal effects between children's cognitive development and stimulation, such effects may exist for other traits or emerge at later time points during children's development.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10519-024-10195-w>.

**Acknowledgements** We thank the participants in the Twins Early Development Study (TEDS) and their families for their ongoing contributions.

**Author Contributions** Olakunle Oginni and Alexandra Starr contributed equally to this work.

**Funding** TEDS is supported by the UK Medical Research Council (MR/V012878/1 and previously MR/M021475/1), with additional support from the US National Institutes of Health (AG046938). This research was funded by the Nuffield Foundation (EDO/44110). OO was supported by the Welsh Clinical Academic Track Fellowship during the writing of this manuscript. SvS was supported by fellowships from the British Academy (2022–2023), the Jacobs Foundation (2022–2027), and the Paris Institute of Advanced Study (2023–2024) during the writing of this manuscript.

**Data Availability** Open Practices Statement: The data for this study can be accessed here: <https://www.teds.ac.uk/researchers/teds-data-access-policy>. The study materials can be accessed here: <https://teds.ac.uk/datadictionary/home.htm>. The analysis code for this study is available on github: <https://github.com/Kaoginni/MRDoC-and-Cross-lagged-model-adjusting-for-cross-time-rA-and-rC/blob/main/Oginni%20C%20Starr%20and%20von%20Stumm%202024.R>. There is a preregistration for this study: available on the OSF: <https://doi.org/10.17605/OSF.IO/D7Y59>.

## Declarations

**Competing Interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Asbury K, Dunn JF, Pike A, Plomin R (2003) Nonshared environmental influences on individual differences in early behavioral development: a monozygotic twin differences study. *Child Dev* 74(3):933–943. <https://doi.org/10.1111/1467-8624.00577>
- Avinun R (2020) The E is in the G: gene–environment–trait correlations and findings from Genome-Wide Association Studies. *Perspect Psychol Sci* 15(1):81–89. <https://doi.org/10.1177/1745691619867107>
- Bailey DH, Duncan GJ, Cunha F, Foorman BR, Yeager DS (2020) Persistence and fade-out of Educational-intervention effects: mechanisms and potential solutions. *Psychol Sci Public Interest* 21(2):55–97. <https://doi.org/10.1177/1529100620915848>
- Bayley N (1993) Bayley Scales of Infant Development. 2nd Ed. San Antonio, TX: Psychological, Corporation
- Blaggan S, Guy A, Boyle EM, Spata E, Manktelow BN, Wolke D, Johnson S (2014) A parent questionnaire for developmental screening in infants born late and moderately Preterm. *Pediatrics* 134(1):e55–e62. <https://doi.org/10.1542/peds.2014-0266>
- Bradley RH (2015) Constructing and adapting causal and formative measures of family settings: the HOME Inventory as Illustration. *J Family Theory Rev* 7(4):381–414. <https://doi.org/10.1111/jftr.12108>
- Breit M, Scherrer V, Tucker-Drob EM, Preckel F (2024) The stability of cognitive abilities: a meta-analytic review of longitudinal studies. *Psychol Bull* 150(4):399–439. <https://doi.org/10.1037/bul0000425>
- Briley DA, Tucker-Drob EM (2013) Explaining the increasing heritability of cognitive ability across development: a meta-analysis of longitudinal twin and adoption studies. *Psychol Sci* 24(9):1704–1713. <https://doi.org/10.1177/0956797613478618>
- Brownell MD, Ekuma O, Nickel NC, Chartier M, Koseva I, Santos RG (2016) A population-based analysis of factors that predict early language and cognitive development. *Early Child Res Q* 35:6–18. <https://doi.org/10.1016/j.ecresq.2015.10.004>
- Burgess S, Thompson SG (2011) Bias in causal estimates from mendelian randomization studies with weak instruments. *Stat Med* 30(11):1312–1323. <https://doi.org/10.1002/sim.4197>
- Burt SA, McGue MATT, Krueger RF, Iacono WG (2005) How are parent–child conflict and childhood externalizing symptoms related over time? Results from a genetically informative cross-lagged study. *Dev Psychopathol* 17(1):145–165. <https://doi.org/10.1017/s095457940505008x>
- Caldwell BM, Bradley RH (1984) Home observation for measurement of the environment, revised edition. University of Arkansas, Little Rock, Arkansas
- Calvin CM, Deary IJ, Webbink D, Smith P, Fernandes C, Lee SH, Luciano M, Visscher PM (2012) Multivariate genetic analyses of cognition and academic achievement from two population samples of 174,000 and 166,000 school children. *Behav Genet* 42(5):699–710. <https://doi.org/10.1007/s10519-012-9549-7>
- d’Apice K, Latham RM, von Stumm S (2019) A naturalistic home observational approach to children’s language, cognition, and behavior. *Dev Psychol* 55(7):1414–1427. <https://doi.org/10.1037/dev0000733>
- Davey Smith G, Ebrahim S (2003) Mendelian randomization’: can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 32(1):1–22. <https://doi.org/10.1093/ije/dyg070>
- Demange PA, Hottenga JJ, Abdellaoui A, Eilertsen EM, Malanchini M, Domingue BW, Armstrong-Carter E, De Zeeuw EL, Rimfeld K, Boomsma DI, Van Bergen E, Breen G, Nivard MG, Cheesman R (2022) Estimating effects of parents’ cognitive and non-cognitive skills on offspring education using polygenic scores. *Nat Commun* 13(1):4801. <https://doi.org/10.1038/s41467-022-32003-x>
- Dick DM (2011) Gene–environment interaction in psychological traits and disorders. *Ann Rev Clin Psychol* 7:383–409
- Dowdall N, Melendez-Torres GJ, Murray L, Gardner F, Hartford L, Cooper PJ (2020) Shared Picture Book Reading interventions for Child Language Development: a systematic review and Meta-analysis. *Child Dev* 91(2). <https://doi.org/10.1111/cdev.13225>
- Enders CK, Bandalos DL (2001) The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Struct Equ Model* 8(3):430–457. [https://doi.org/10.1207/S15328007SEM0803\\_5](https://doi.org/10.1207/S15328007SEM0803_5)
- Erbeli F, van Bergen E, Hart SA (2020) Unraveling the relation between reading comprehension and print exposure. *Child Dev* 91(5):1548–1562. <https://doi.org/10.1111/cdev.13339>
- Feinstein L, Bynner J (2004) The importance of cognitive development in middle childhood for adulthood socioeconomic status, mental health, and problem behavior. *Child Dev* 75(5):1329–1339. <https://doi.org/10.1111/j.1467-8624.2004.00743.x>
- Fenson L, Marchman VA, Thal DJ, Dale PS, Reznick JS, Bates E (2007) MacArthur-Bates communicative development inventories: User’s guide and technical manual (2nd ed.). Brookes



- Gillespie NA, Zhu G, Neale MC, Heath AC, Martin NG (2003) Direction of causation modeling between cross-sectional measures of parenting and psychological distress in female twins. *Behav Genet* 33(4):383–396. <https://doi.org/10.1023/A:1025365325016>
- Grotzinger AD, Cheung AK, Patterson MW, Harden KP, Tucker-Drob EM (2019) Genetic and Environmental Links between General Factors of Psychopathology and cognitive ability in early childhood. *Clin Psychol Sci* 7(3):430–444. <https://doi.org/10.1177/2167702618820018>
- Hart B, Risley TR (1995) Meaningful differences in the everyday experience of young American children. Paul H Brookes Publishing
- Heath AC, Kessler RC, Neale MC, Hewitt JK, Eaves LJ, Kendler KS (1993) Testing hypotheses about direction of causation using cross-sectional family data. *Behav Genet* 23(1):29–50. <https://doi.org/10.1007/BF01067552>
- Howard SJ, Powell T, Vasseleu E, Johnstone S, Melhuish E (2017) Enhancing preschoolers' executive functions through embedding cognitive activities in Shared Book Reading. *Educational Psychol Rev* 29(1):153–174. <https://doi.org/10.1007/s10648-016-9364-4>
- Kohler HP, Behrman JR, Schnitker J (2011) Social science methods for twins data: integrating causality, endowments, and heritability. *Biodemography Social Biology* 57(1):88–141. <https://doi.org/10.1080/02F19485565.2011.580619>
- Krapohl E, Hannigan LJ, Pingault JB, Patel H, Kadeva N, Curtis C, Breen G, Newhouse SJ, Eley TC, O'Reilly PF, Plomin R (2017) Widespread covariation of early environmental exposures and trait-associated polygenic variation. *PNAS* 114(44):11727–11732. <https://doi.org/10.1073/pnas.1707178114>
- Lee J, Wedow R, Okbay A, Team aR, Social Science Genetic Association Consortium (2018) Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet* 50(8):1112–1121. <https://doi.org/10.1007/BF01067552>. COGENT (Cognitive Genomics Consortium)
- Martin AJ, Darlow BA, Salt A, Hague W, Sebastian L, McNeill N, Tarnow-Mordi W (2013) Performance of the parent report of children's abilities-revised (PARCA-R) versus the Bayley scales of Infant Development III. *Arch Dis Child* 0:1–4. <https://doi.org/10.1136/archdischild-2012-303288>
- McAdams TA, Rijdsdijk FV, Zavos HM, Pingault JB (2021) Twins and causal inference: leveraging nature's experiment. *Cold Spring Harbor Perspect Med* 11(6):a039552. <https://doi.org/10.1101/cshperspect.a039552>
- McCarthy D (1972) McCarthy Scales of Children's abilities. The Psychological Corporation, New York
- Minică CC, Dolan CV, Boomsma DI, de Geus E, Neale MC (2018) Extending causality tests with genetic instruments: an integration of mendelian randomization with the classical twin design. *Behav Genet* 48(4):337–349. <https://doi.org/10.1007/s10519-018-9904-4>
- Mullen G (2017) More than words: using nursery rhymes and songs to support domains of child development. *J Child Stud* 42(2):42–53. <https://doi.org/10.18357/jcs.v42i2.17841>
- Neale MC, Cardon LR (1992) Methodology for Genetic Studies of Twins and families. Springer Netherlands. <https://doi.org/10.1007/978-94-015-8018-2>
- Neale MC, Hunter MD, Pritikin JN, Zahery M, Brick TR, Kirkpatrick RM, Estabrook R, Bates TC, Maes HH, Boker SM (2016) OpenMx 2.0: extended structural equation and statistical modeling. *Psychometrika* 81(2):535–549. <https://doi.org/10.1007/s11336-014-9435-8>
- NICE NI (2017) for C. E. Developmental follow-up of children and young people born preterm, NICE guideline [NG72]. NICE. <https://www.nice.org.uk/guidance/ng72/chapter/Recommendations#providing-enhanced-developmental-support>
- Noble C, Sala G, Peter M, Lingwood J, Rowland C, Gobet F, Pine J (2019) The impact of shared book reading on children's language skills: a meta-analysis. *Educational Res Rev* 28:100290. <https://doi.org/10.1016/j.edurev.2019.100290>
- Oliver BR, Plomin R (2007) Twins' early Development Study (TEDS): a Multivariate, Longitudinal Genetic Investigation of Language, cognition and behavior problems from Childhood through Adolescence. *Twin Res Hum Genet* 10(1):96–105. <https://doi.org/10.1375/twin.10.1.96>
- Oliver B, Dale PS, Saudino KJ, Petrill SA, Pike A, Plomin R (2002) The validity of a parent-based assessment of cognitive abilities in 3-year-olds. *Early Child Dev Care* 172(4):337–348. <https://doi.org/10.1080/03004430212713>
- Petrill SA, Lipton PA, Hewitt JK, Plomin R, Cherny SS, Corley R, DeFries JC (2004) Genetic and environmental contributions to General Cognitive ability through the first 16 years of life. *Dev Psychol* 40(5):805–812. <https://doi.org/10.1037/0012-1649.40.5.805>
- Plomin R (1995) Genetics and children's experiences in the family. *J Child Psychol Psychiatry* 36(1):33–68. <https://doi.org/10.1111/j.1469-7610.1995.tb01655.x>
- Plomin R (2011) Commentary: why are children in the same family so different? Non-shared environment three decades later. *Int J Epidemiol* 40(3):582–592. <https://doi.org/10.1093/ije/dyq144>
- Plomin R, DeFries JC, Loehlin JC (1977) Genotype-environment interaction and correlation in the analysis of human behavior. *Psychol Bull* 84:309–322. <https://doi.org/10.1037/0033-2909.84.2.309>
- Polderman TJ, Benyamin B, De Leeuw CA, Sullivan PF, Van Bochoven A, Visscher PM, Posthuma D (2015) Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat Genet* 47(7):702. <https://doi.org/10.1038/ng.3285>
- Price TS (2002) Genetic and environmental influences on verbal and nonverbal domains of cognitive ability and low cognitive ability: A study of 2-, 3-, and 4-year-old twins. Ph.D Dissertation
- Price TS, Jaffee SR (2008) Effects of the family environment: gene-environment interaction and passive gene-environment correlation. *Dev Psychol* 44(2):305–315. <https://doi.org/10.1037/0012-1649.44.2.305>
- Protzko J (2015) The environment in raising early intelligence: a meta-analysis of the fadeout effect. *Intelligence* 53:202–210. <https://doi.org/10.1016/j.intell.2015.10.006>
- Rasmussen SHR, Ludeke S, Hjelmberg JV (2019) A major limitation of the direction of causation model: non-shared environmental confounding. *Twin Res Hum Genet* 22(1):14–26. <https://doi.org/10.1017/thg.2018.67>
- Rijdsdijk FV, Sham PC (2002) Analytic approaches to twin data using structural equation models. *Brief Bioinform* 3(2):119–133. <https://doi.org/10.1093/bib/3.2.119>
- Rimfeld K, Malanchini M, Spargo T, Spickernell G, Selzam S, McMillan A, Plomin R (2019) Twins early development study: a genetically sensitive investigation into behavioral and cognitive development from infancy to emerging adulthood. *Twin Res Hum Genet* 22(6):508–513. <https://doi.org/10.1017/thg.2019.56>
- Sanderson E et al (2022) Mendelian randomization. *Nature Review Methods Primers* 2, 7 (2022). <https://doi.org/10.1038/s43586-022-00099-6>
- Saudino KJ, Dale PS, Oliver B, Petrill SA, Richardson V Rutter, M, Simonoff E, Stevenson J, Plomin R (1998) The validity of parent-based assessment of the cognitive abilities Of 2-year-olds. *Br J of Developmental Psychology* 16, 349–363. <https://doi.org/10.1111/j.2044-835X.1998.tb00757.x>
- Starr A, Ruks M, Weigel L, Riemann R (2023) What drives the association between home chaos and school grades over time? A biometric cross-lagged panel approach. *Learn Individual Differences* 104:102287. <https://doi.org/10.1016/j.lindif.2023.102287>
- Tachibana Y, Fukushima A, Saito H, Yoneyama S, Ushida K, Yoneyama S, Kawashima R (2012) A new mother-child play



- activity program to decrease parenting stress and improve child cognitive abilities: a Cluster Randomized Controlled Trial. *PLoS ONE* 7(7):e38238. <https://doi.org/10.1371/journal.pone.0038238>
- Takacs ZK, Kassai R (2019) The efficacy of different interventions to foster children's executive function skills: a series of meta-analyses. *Psychol Bull* 145(7):653–697. <https://doi.org/10.1037/bul0000195>
- Tick B, Colvert E, McEwen F, Stewart C, Woodhouse E, Gillan N, Hallett V, Lietz S, Garnett T, Simonoff E (2016) Autism Spectrum disorders and other mental health problems: exploring etiological overlaps and phenotypic causal associations. *J Am Acad Child Adolesc Psychiatry* 55(2):106–113. <https://doi.org/10.1016/j.jaac.2015.11.013>
- Tucker-Drob EM, Harden KP (2012) Early childhood cognitive development and parental cognitive stimulation: evidence for reciprocal gene–environment transactions. *Dev Sci* 15(2):250–259. <https://doi.org/10.1111/j.1467-7687.2011.01121.x>
- Turkheimer E, Waldron M (2000) Nonshared environment: a theoretical, methodological, and quantitative review. *Psychol Bull* 126(1):78–108. <https://doi.org/10.1037/0033-2909.126.1.78>
- van Bergen E, Snowling MJ, de Zeeuw EL, van Beijsterveldt C, Dolan CV, Boomsma DI (2018) Why do children read more? The influence of reading ability on voluntary reading practices. *J Child Psychol Psychiatry* 59(11):1205–1214. <https://doi.org/10.1111/jcpp.12910>
- Verhulst B, Hatemi PK (2013) Gene–environment interplay in twin models. *Political Anal* 21(3):368–389. <https://doi.org/10.1093/pan/mpt005>
- von Stumm S, Rimfeld K, Dale PS, Plomin R (2020) Preschool verbal and nonverbal ability mediate the association between socioeconomic status and school performance. *Child Dev* 91(3):705–714. <https://doi.org/10.1111/cdev.13364>
- von Stumm S, Kandaswamy R, Maxwell J (2023) Gene–environment interplay in early life cognitive development. *Intelligence* 98:101748. <https://doi.org/10.1016/j.intell.2023.101748>
- Wertz J, Belsky J, Moffitt TE, Belsky DW, Harrington H, Avinun R, Poulton R, Ramrakha S, Caspi A (2019) Genetics of nurture: a test of the hypothesis that parents' genetics predict their observed caregiving. *Dev Psychol* 55(7):1461–1472. <https://doi.org/10.1037/dev0000709>
- Xie QW, Chan CH, Ji Q, Chan CL (2018) Psychosocial effects of parent–child book reading Interventions: a meta-analysis. *Pediatrics*, 141(4), e20172675. <https://doi.org/10.1542/peds.2017-2675>
- Xiong X, Deng L, Li H (2020) Is winning at the start important: early childhood family cognitive stimulation and child development. *Child Youth Serv Rev* 118:105431. <https://doi.org/10.1016/j.childyouth.2020.105431>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.