Child Development, 2026, 00(00), 1–13
https://doi.org/10.1093/chidev/aacag097
Published: June 2, 2026
Empirical Article

Society for Research in Child Development
OXFORD

The role of mothers' speech for children's education beyond the early years

Anna Brown[1,*], Avshalom Caspi[2,3,4,5], Helen L. Fisher[2,6], Terrie E. Moffitt[2,3,4,5], Jasmin Wertz[7], and Sophie von Stumm[8]

[1]School of Psychology, University of Northampton, Northampton, United Kingdom
[2]Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom
[3]Department of Psychology and Neuroscience, Duke University, Durham, NC, United States
[4]Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC, United States
[5]PROMENTA, Department of Psychology, University of Oslo, Oslo, Norway
[6]ESRC Centre for Society and Mental Health, King's College London, London, United Kingdom
[7]Department of Psychology, University of Edinburgh, Edinburgh, United Kingdom
[8]Department of Education, University of York, York, United Kingdom
*Corresponding author: School of Psychology, University of Northampton, University Drive, Northampton NN1 5PH, United Kingdom.
Email: anna.brown@northampton.ac.uk

Abstract

We tested the extent to which mothers' speech contributes to the transmission of family background inequality in education. In 894 families (93.1% White), representative of the full range of Britain's socioeconomic conditions, we quantified mothers' vocabulary sophistication, lexical diversity, and grammatical complexity from 10-min-long audio-recorded interviews. Mothers' vocabulary sophistication significantly predicted children's (49% males) cognition, literacy, and educational achievement from ages 5 to 12 years, accounting for 2%–5% of the variance. After adjusting for mothers' education and household income, these effects reduced to 1% and 2% or became nonsignificant. Our findings suggest vocabulary sophistication contributes only modestly to the transmission of family background inequality in education.

Keywords education, mothers' speech, word gap

Lay summary

Research has shown that the way mothers talk to their children before they start school may play a role in passing on educational inequalities. The longer-term impact of mothers' speech beyond the early years is not well understood. Using interviews with 894 mothers from a wide range of socioeconomic backgrounds, we examined whether differences in mothers' vocabulary and grammar are linked to children's outcomes at school. Children whose mothers used more sophisticated vocabulary showed small advantages in cognitive ability and reading between ages 5 and 10. However, when mothers' education and household income were taken into account, most of these links became weaker or disappeared. Overall, mothers' education and household income had a stronger long-term influence on children's educational outcomes.

Doing well in school is the first step to success in meritocratic societies that allocate resources and power according to demonstrated achievement. A child's performance in school predicts the number, level, and prestige of the educational qualifications they are likely to achieve over their lifetime (Duncan et al., 2007; Entwisle et al., 2005; Starr et al., 2024). These educational qualifications, in turn, inform the child's later access to economic, social, and cultural resources (Bourdieu, 1986), their position in society (Carlson & McChesney, 2015), well-being (Noble et al., 2021), and health (Furnée et al., 2008). Yet, children are not equally equipped for doing well in school: children from families with fewer socioeconomic resources tend to lag in school readiness, and they achieve lower grades in primary and secondary education than children from well-resourced families (Magnuson et al., 2016; von Stumm et al., 2022a).

Prior studies identified the speech children hear from their mothers as a key pathway through which family background inequality in education becomes transmitted (Hoff, 2013; Rowe, 2018; Schwab & Lew-Williams, 2016). Child-directed speech supports early language development through well-documented

Handling Editor: Jessica Sullivan
Editor: Shauna Cooper
Received: May 13, 2025. **Revised:** March 12, 2026. **Accepted:** April 14, 2026

mechanisms, such as contingency, diversity of vocabulary, and decontextualized language use (Rowe & Snow, 2020). These mechanisms can transmit family background inequality as mothers from under-resourced families speak, on average, less frequently to their children (Rowe, 2008), and they use simpler vocabulary and syntax than mothers from well-resourced families (Huttenlocher et al., 2010; Rowe, 2012). The quantity and quality of child-directed speech support children's vocabulary and syntactic growth, which are the foundation for later literacy skills and academic achievement (Durham et al, 2007; Rowe, 2012). It follows that children's early life language experiences and their language development are both conditioned on their family background in ways that can foster inequalities in education (Hoff, 2013; Pace et al., 2017; Rowe, 2018; Schwab & Lew-Williams, 2016).

Previous research focused on child-directed speech but paid less attention to the role that mothers' adult-directed speech plays in children's development. This omission is striking for two reasons. First, children begin to overhear mothers' adult-directed speech in utero, and this exposure permeates across contexts (e.g., home versus school). A vast body of empirical studies demonstrated that overhearing produces robust learning in children (Akhtar, 2005; Girouard-Hallam & Norris, 2024; Sperry et al., 2019). Exposure to overheard adult-directed speech may complement ongoing contingent, developmentally tailored interactions, such as child-directed speech. Both forms of speech likely exert long-term influence on children's development, albeit to different degrees and potentially through distinct pathways. Second, adult-directed speech may capture a broader language socialization pathway, whereby mothers' adult-directed speech models academic registers, discourse norms, and cognitively demanding language practices that align with school-based expectations (Ochs & Schieffelin, 2017; Schleppegrell, 2004). It therefore seems plausible that mothers' adult-directed speech contributes not only to children's language development but also to their cognition, literacy, and educational achievement.

Previous studies' narrow focus on the influence of child-directed speech may have reinforced the idea that the transmission of family background inequality can be explained by a single pathway (i.e., mothers' child-directed speech). However, this notion risks oversimplifying the role of mothers' speech for children's differences in development (e.g., Sperry et al., 2019). To address this gap, we tested here the long-term prediction of children's cognition, literacy, and educational achievement at age 5–12 years from mothers' adult-directed speech, including their vocabulary sophistication, lexical diversity, and grammatical complexity. We use the term "prediction" in this article to refer to the temporal order in which the assessment of one variable precedes the assessment of another.

Mothers' speech may contribute to children's differences in education via two principal mechanisms that likely intersect and operate along a social gradient (Hart & Risley, 1995; Lippi-Green, 2011). For one, mothers' speech frames children's opportunities to acquire vocabulary and grammar, which influence their development in domains that thrive through language exposure, such as reasoning (Mercer et al., 1999), literacy (Duff et al., 2015), and school performance (Pace et al., 2019). For example, children who hear a wide range of sophisticated vocabulary from their mothers show, on average, faster word recognition and deeper text comprehension than children who experience less vocabulary sophistication (Verhoeven et al., 2011). Word recognition and text comprehension, in turn, enhance children's literacy (Verhoeven et al., 2011).

For the other, mothers' speech can familiarize children with the formal academic register that is prevalent in educational contexts, in line with Bourdieu's (1986) theory of linguistic capital (Lareau & Weininger, 2003; Schleppegrell, 2004). Registers refer to patterned ways of language use that vary across social contexts and communication activities (see Halliday & Hasan, 1989). Here, the academic register describes a specialized form of language that is valued in academic settings and characterized by clusters of complex, co-occurring lexical and grammatical features. Such lexical features include, for example, subject-specific and low-frequency terminology, while the grammatical features include complex sentence structures, nominalizations, and passive constructions (Fang et al., 2006; Hadley et al., 2022). Mothers whose speech aligns with the academic register model a form of language use for their children that resembles the language used in classrooms and academic texts (i.e., language socialization). Children who are familiar with the academic register when they start formal schooling are at an advantage because they can readily access the curriculum and engage with the teachers. However, children raised without many opportunities to learn the linguistic codes that govern education settings may—unwittingly—communicate in ways that lead to being stereotyped and discriminated against by teachers (Kircher & Fox, 2021; Spencer et al., 2013).

Current study

Our study takes a confirmatory approach to testing the role mothers' speech plays in the transmission of family background inequality in education, focusing on three preregistered hypotheses. First, we tested the hypothesis that characteristics of mothers' speech, including vocabulary sophistication, lexical diversity, and grammatical complexity, would significantly, though modestly predict children's long-term cognition, literacy, and educational achievement, including verbal and nonverbal skills at age 5 years, reading ability at ages 7 and 10 years, and academic performance in English and Maths at ages 7, 10, and 12 years. We also tested whether these predictions are independent of SES or attenuated when SES is accounted for. Our second hypothesis was that markers of mothers' SES, including occupation, educational attainment, and household income, were associated with the characteristics of their speech in the way that mothers of higher SES spoke with higher vocabulary sophistication, lexical diversity, and grammatical complexity. Finding such positive correlations would suggest that mothers' speech differences are shaped by a "social gradient", whereby mothers with fewer socioeconomic resources are also at a greater risk of using speech that is less beneficial for their children's development (cf. Marmot, 2015; Mulder & Hulstijn, 2011). Our third hypothesis was that mothers' speech mediates the association between mothers' SES and children's education-related outcomes. To test our analyses, we conducted preregistered regression and mediation analyses that together clarify the extent to which mothers' speech contributes to the intergenerational transmission of educational advantage, both directly and indirectly via SES.

We capitalized on rich, naturalistic speech data collected from a socioeconomically diverse sample of $N = 894$ mothers when their

twin children were 5 years old and starting primary school (Caspi et al., 2004). Data came from E-Risk, a UK population-representative cohort study of families who had twins born from 1994 to 1995 in England and Wales (Moffitt & E-Risk Study Team, 2002). The mother–child trios in E-Risk were identified via the Office for National Statistics' multibirth register (Moffitt & E-Risk Study Team, 2002) and represented the full range of Britain's socioeconomic conditions, with an ethnic composition comparable to that of England and Wales, at the time (e.g., 94.1% White families in England and Wales; 93.1% White families in E-Risk). E-Risk families' residential addresses match the UK Index of Multiple Deprivation distribution: 25% of the families lived in Britain's wealthiest areas and 25% in the poorest areas (Odgers et al., 2012). Because E-Risk has suffered minimal attrition since its inception (∼5%), in contrast to many other longitudinal studies (Gustavson et al., 2012), the sample's representativeness remained intact over time.

Mothers freely described their parenting experiences and the twins' characteristics in interviews conducted in person by trained health researchers (i.e., adult interviewers; average interview length ∼22 min; Caspi et al., 2004). From 10-min-long excerpts of these interviews' transcripts, we quantified mothers' lexicon in terms of vocabulary sophistication (% of rare words used in speech; Weizman & Snow, 2001) and lexical diversity (range of word types used in speech; McCarthy & Jarvis, 2010). We also quantified mothers' grammatical complexity from the mean length of utterance, verb usage (number of verbs per utterance; Hoff-Ginsberg, 1990), and syntactic complexity (proportion of grammatical relations reflecting complex syntactic structures; Mueller et al., 2018). Children's verbal and nonverbal cognitive abilities at age 5 years and literacy skills (i.e., reading ability) at age 7 and 10 years were assessed with standardized psychometric tests. Children's performance in the core school subjects of English and Mathematics at ages 7, 10, and 12 years was rated by teachers relative to that of other pupils at the same age.

Materials and methods

Sample

Data came from the Environmental Risk (E-Risk) Longitudinal Twin Study, which tracks the development of 2,232 British twin children drawn from a larger birth register of twins born in England and Wales in 1994 and 1995 (Trouton et al., 2002). Full details about the sample are reported elsewhere (Moffitt & E-Risk Study Team, 2002) and in Supplementary Note S4. To summarize, the E-Risk sample was constructed in 1999 to 2000 when 1,116 families (93% of those eligible and of whom 90.4% were White) with same-sex 5-year-old twins participated in home-visit assessments. This sample comprised 56% monozygotic and 44% dizygotic twin pairs; sex was evenly distributed within zygosity (49% male). Families were recruited to represent the UK population of families with newborns in the 1990s, on the basis of residential location throughout England and Wales and mothers' age (i.e., older mothers having twins via assisted reproduction were under-selected, and teenage mothers with twins were over-selected). This resulted in a sample of families that were representative of the distribution of neighborhood socio-economic status across the UK population (Odgers et al., 2012).

Follow-up home visits were conducted when children were aged 7, 10, 12, and 18 years (participation rates were 98%, 96%, 96%, and 93%, respectively). Here, we use data collected up to age 12, when 1,073 families participated. Home visits at ages 5, 7, 10, and 12 years included assessments with participants and their mother (or primary caregiver); a different interviewer assessed each participant in a twin pair. With parents' permission, questionnaires were posted to the children's teachers, who returned questionnaires for 94% of children at age 5, 91% of the 2,232 E-Risk children (93% of those followed up) at age 7, 86.3% of the 2,232 E-Risk children (90.1% of those followed up) at age 10, and 80% of the 2,232 E-Risk children at age 12 (83% of those followed up). The Joint South London and Maudsley and Institute of Psychiatry Research Ethics Committee approved each study phase. Parents gave informed consent, and twins gave assent between 5 and 12 years and then informed consent at age 18.

Mothers' speech corpus

When their twin children were 5 years old (recorded approx. 2,000), mothers freely described their parenting experiences and the twins' characteristics in interviews conducted in person by trained health researchers (Caspi et al., 2004). Mothers were encouraged to talk freely with few interruptions. However, if a mother found this difficult, the interviewer could aid her with a series of semi-structured probes, such as "In what ways would you like [child] to be different?". Interviews about each twin were separated in time by approximately 90 min. All interviews were audiotaped with the mothers' consent. Interview data were missing for 6% of the sample because some mothers did not wish to be audiotaped, because of technical problems with the tape, or because of loss to follow-up. The audio-taped interviews were transcribed, and 10-minute samples from each interview (to ensure comparability given the varying interview lengths) were coded for automatic analyses using the Computer Language Analysis (CLAN) program (MacWhinney, 2000). A further 1% of interviews were excluded at the transcription stage due to poor audio. 2% of families were then excluded because a language other than English was spoken at home. Our final analysis sample included all mothers and their twins with sufficient available speech data (i.e., 1,788 children from 894 families, of whom 93.1% were White).

Mothers' speech characteristics

From the transcripts of mothers' speech, we extracted quantitative markers of lexicon and grammar using the FREQ, VOCD, and MLU programs within CLAN (MacWhinney, 2000). For lexical markers, we extracted mothers' vocabulary sophistication and lexical diversity. For grammar markers, we extracted mothers' mean length of utterance (morpheme count), grammatical complexity index, and verb usage within utterances. When extracting grammar markers, partially unintelligible utterances were excluded. For descriptives and correlations of lexical and grammatical markers, see Supporting Information (Table S1 and Figure S1).

Vocabulary sophistication

Vocabulary sophistication was the percentage of rare words mothers used in their respective speech transcripts. From the

transcripts, we removed all words on the Dale-Chall word list (Chall & Dale, 1995), their inflected forms, all nondictionary words, and any derivations of family names. The Dale-Chall word list includes the 3,000 most common words known by fourth graders (Chall et al., 1995). We converted the words on the Dale-Chall list to UK spelling to align with the linguistic context of our sample. After filtering words, approximately 10% of transcripts were screened and manually checked by the first author and calibrated by trained research assistants to ensure the accuracy of this filtering process. Any remaining words in the transcripts were considered rare and sophisticated word types (Weizman & Snow, 2001). The percentage of mothers' rare word types relative to their total number of word types in the transcript was then calculated.

Lexical diversity

Lexical diversity was measured by computing D scores, which index the probability of introducing a new word as the speech sample progresses (MacWhinney, 2000). D-scores are highly robust against sample size effects and thus constitute reliable markers of lexical diversity over other methods, such as Type Token Ratio (McKee et al., 2000)

Grammatical complexity

Grammatical complexity was inferred from the grammatical complexity index (GCI; Mueller et al., 2018), verb index, and mean length of utterance (Hoff-Ginsberg, 1990).

GCI was computed by calculating mothers' number of grammatical relations that marked syntactic embeddings divided by their total number of grammatical relations (Mueller et al., 2018); the higher the index, the more complex grammatical and syntactic structures used in speech. The verb index was the mother's total number of verbs divided by the total number of utterances. The higher the average number of verbs per utterance, the more complex the syntax. MLU was calculated by dividing the mothers' total number of morphemes by their total number of utterances produced. The higher the MLU, the more complex and longer mothers' utterances were. Any words classed as communicators, such as "oh" or "ah", were excluded from this analysis.

Preliminary analyses were carried out to validate the reliability of the characteristics of mothers' speech. Our lexical markers (vocabulary sophistication and lexical diversity) were positively intercorrelated at $r = .34$, and our grammar markers (MLU, grammatical complexity, and verb usage) correlated between .79 and .91. Because the grammar markers' correlations were above .60, we computed a grammar composite score by summing the *z*-scores for grammatical complexity, MLU, and verb usage, in line with our preregistration. Mothers' reading scores from the Wide-Range Achievement Test (WRAT-3, Snelbaker et al., 2001) were positively correlated with vocabulary sophistication ($r = .40$), lexical diversity ($r = .17$), and grammatical score ($r = .12$). Mothers' complexity of speech and WRAT-3 scores have been shown elsewhere to be highly correlated (Golden & Bennett, 2024), which aligns with our findings here.

SES indicators

Mothers' educational attainment

The highest educational qualification of mothers when the children were aged 5 were coded into 5 groups (1= No Qualifications; 2 = Certificate of Secondary Education (2–5)/General Certificate of Secondary Education (D-G), i.e., below average school leaving qualification; 3= Certificate of Secondary Education (1)/O Level (A-C)/General Certificate of Secondary Education (A-C), i.e., above average school qualification; 4= Advanced Level/Scholarship Level; 5 = Higher National Certificate/Degree i.e., College Level Qualification; 6= Higher National Diploma/Postgraduate degree).

Mothers' occupation

The mothers' current (or most recent) occupation when the children were aged 5 was coded using the Office of Population Censuses et al. (1995) Standard Occupational Classification. Occupational groups are arranged into six social classes (1 = professional occupations; 2 = managerial and technical occupations; 3 = skilled occupations (nonmanual); 4 = skilled occupations (manual); 5 = partly skilled occupations; 6 = unskilled occupations).

Household income

Mothers reported the household's total income from all sources before tax in the previous 12 months when the children were aged 5. Families were then coded into 16 income brackets, the lowest bracket representing an annual income of £10,000 or less and the highest representing an annual income of £40,000 or more.

Cognition and literacy

We tested the extent to which mothers' speech predicted children's cognition and literacy concurrently and prospectively at the ages of 5, 7, and 10 years, assessed by standardized psychometric tests (Torgesen et al., 1999; Wechsler, 1989), which in turn forecast their educational achievement (Pace et al., 2019). Reliability estimates exceed .90 for all measures in the E-Risk cohort (Rutter et al., 2004).

Verbal and nonverbal ability

A short form of the Wechsler Preschool and Primary Scale of Intelligence-Revised (Wechsler, 1989) was used to assess the cognitive ability of young children at age 5 years. Two subtests were administered: a score on the Vocabulary Subtest, taken as a measure of verbal ability and a score on the Block Design Subtest, taken as a measure of nonverbal ability. All raw scores were converted to age-based standard scores. The Vocabulary Subtest and the Block Design Subtest scores were treated as independent outcomes that represented verbal and nonverbal ability, respectively.

Reading score

Children completed the Sight Word Efficiency (SWE) subtest of the Test of Word Reading Efficiency (TOWRE), which measures a child's ability to pronounce printed words accurately and fluently (Torgesen et al., 1999). The SWE was administered at ages 7 and 10 years to assess the number of real printed words that children can identify accurately within 45 s. All raw scores were converted to age-based standard scores. The Sight Word Efficiency subtest has high internal consistency and developmental sensitivity (Tarar et al., 2015), is associated with oral language development (Killingly et al., 2025), and correlates strongly with comprehension measures (Babayiğit, 2015).

Educational achievement

Teachers rated the children's performance in English and Mathematics at ages 7, 10, and 12 using a 5-point Likert scale (1 = Far Below, 5 = Far Above), comparing each child to a typical peer of the same age. Prior research has demonstrated that

teacher assessments are robust and valid measures of academic achievement in UK samples. For example, teacher ratings correlated strongly at ∼.70 with standardized test scores across core subjects and school years (i.e., ages 7–14 years) in a population-representative cohort sample of up to 6,300 British children (Rimfeld et al., 2019). In E-Risk, teacher ratings of English and Mathematics were highly intercorrelated ($r = .81–.84$), a finding that converges with those from analyses of school grades in comparable samples (e.g., von Stumm et al., 2022a, 2022b).

Initially, we fitted separate latent growth curve models to both subjects' grades. However, because of the high intercorrelations between English and Maths grades, and because latent growth factors were highly similar across subjects, we report findings from the models that included grades across English and Maths (i.e., summed scores) in the main manuscript. We report the outcomes of the subject-specific models in the Supplementary Materials (see Table S8).

Statistical analysis

All models were fitted using the R package Lavaan (Rosseel, 2012). To account for the nonindependence of observations (i.e., twins), standard errors were clustered at the family level, and mothers' age was included as a covariate in all models. E-Risk incurred very low attrition over time (∼5%) with data missing randomly; we used Full Information Maximum Likelihood (FIML) in all models (Graham, 2009). We fitted hierarchical linear regression models to test whether mothers' speech was associated with their SES and predicted their children's cognition and literacy. An independent model was run for each outcome at each age (verbal ability at age 5 years, nonverbal ability at age 5 years, reading score at age 7 years, and reading score at age 10 years).

We fitted latent growth curve models to test whether mothers' speech was associated with their children's educational achievement. Using latent growth curve models allowed us to discern associations between mothers' speech and children's variance in educational achievement that was stable throughout the school years (intercept) versus variance in educational achievement that reflected systematic gains and losses in school performance over time (slope; see details in Supporting Information). This approach enabled testing whether mothers' lexicon and grammar predicted differences in children's educational achievement that were stable across the school years or predicted dynamic changes in educational achievement over time (Preacher, 2018).

Finally, to discern the effects of mothers' speech from those of SES, we conducted mediation models to estimate the effect size of mothers' speech as a mediator of the SES–achievement association, using language indicators that were significant predictors in prior regression models. A random twin was selected from each pair to account for the nonindependence of observations within twin families ($N = 894$).

We conducted mediation analyses to estimate the indirect and direct effects of mothers' speech in the association between SES and children's education-related outcomes. Our approach was motivated by the conceptual logic of mediation (Baron & Kenny, 1986; MacKinnon et al., 2007) but was not contingent on the significance-testing framework proposed in those models. Instead, mediation was used to quantify the relative contribution of mothers' speech to the SES–outcome association, providing an effect-size estimate of the indirect pathway. Mediators were identified based on prior regression analyses that established significant associations with both SES and children's outcomes. Indirect effects were estimated using bias-corrected bootstrapped confidence intervals (95% CI, 1,000 resamples; Li, 2011), and mediation percentages were calculated to compare the proportion of variance explained by indirect and direct effects (Hayes, 2009; Zhao et al., 2010).

Results

We present our findings for children's cognition and literacy separately from those pertaining to children's educational achievement, reflecting the different modeling approaches that we applied to these outcomes (i.e., regressions versus latent growth curve models).

Predicting children's cognition and literacy from mothers' speech

We fitted independent models for each measure at each age with mothers' speech characteristics as predictors.

Mothers' vocabulary sophistication was significantly and positively associated with children's cognition and literacy (see Tables S2–4), accounting for 2%–5% of the variance across ages and measures (Figure 1A). Neither mothers' lexical diversity nor grammar significantly predicted children's cognition and literacy (Figure 1A).

Predicting children's educational achievement from mothers' speech

We next tested the extent to which mothers' speech predicted children's school performance in English and Mathematics, as rated by teachers when the children were aged 7, 10, and 12 years.

We specified mothers' speech characteristics (i.e., vocabulary sophistication, lexical diversity, and grammatical complexity) as predictors of the two latent growth factors (i.e., intercept and slope). Mothers' vocabulary sophistication was significantly associated with children's stable variance in educational achievement (intercept), accounting for 2% of the variance, but not with gains and losses in educational achievement over time (slope; Table S5). Thus, mothers' vocabulary sophistication predicted children's achievement differences that persisted throughout school with a small effect size but did not forecast children's improvements or worsening in school performance over time relative to where they started. Mothers' lexical diversity and grammatical complexity were associated neither with the intercept nor the slope of children's educational achievement (Table S5). Akin to our findings on children's cognition and literacy, these results suggest that mothers' vocabulary sophistication predicts children's differences in educational achievement. Yet, mothers' vocabulary sophistication was not associated with changes in achievement trajectories that occur over time.

Mothers' SES and characteristics of their speech

To explore the role mothers' speech plays in the transmission of family background inequality, we first tested whether mothers'

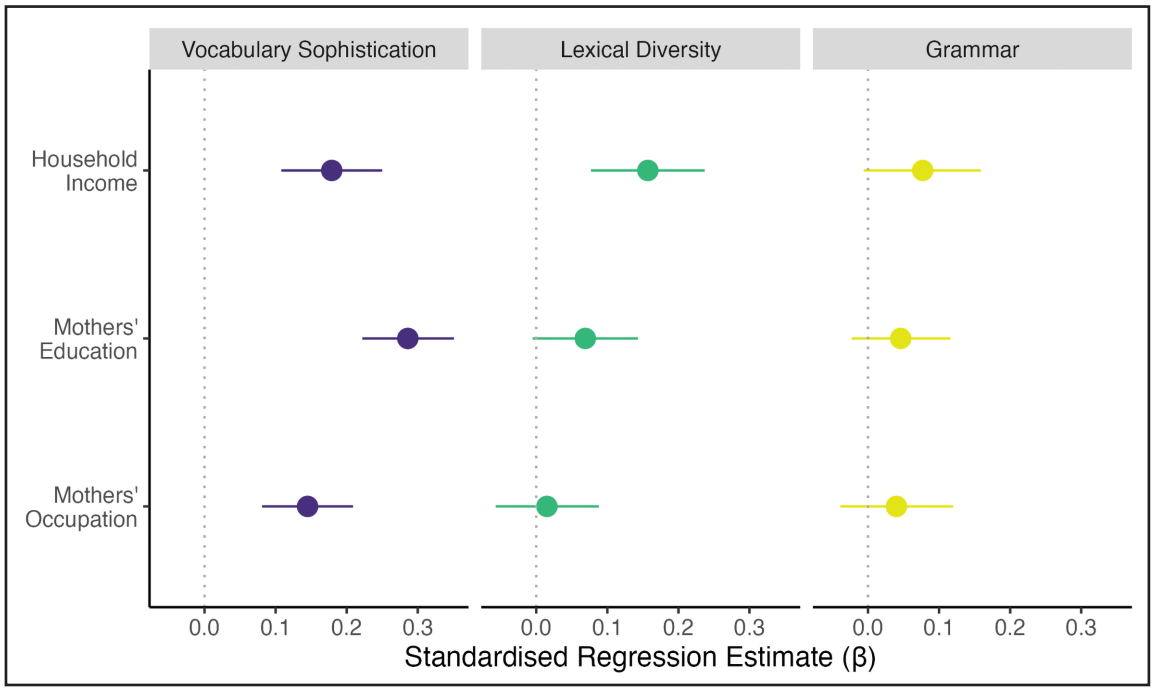

Figure 1 Predicting children's cognition and literacy from mothers' speech and SES. Note. Plot (A) (left) shows the associations between mothers' speech and children's cognition and literacy; Plot (B) (middle) shows associations between mothers' socioeconomic status (SES) and children's cognition and literacy; Plot (C) (right) shows the association between mothers' speech and children's cognition and literacy after adjusting for SES. Dots represent standardized (beta) regression estimates, and lines represent 95% confidence intervals. All models were adjusted for mothers' age.

SES, including their educational attainment, occupation, and household income, was associated with their speech. We fitted independent regression models for each characteristic of mothers' speech (i.e., vocabulary sophistication, lexical diversity, and grammatical complexity) to be predicted by mothers' SES indicators.

Mothers' SES indicators accounted for 28% of the variance in vocabulary sophistication and 6% of the variance in lexical diversity; they were not significantly associated with grammatical complexity (Figure 2, Table S6). These results suggest that only measures of mothers' lexicon (i.e., vocabulary sophistication and lexical diversity) vary as a function of their SES.

Mothers' educational attainment was the strongest independent correlate of vocabulary sophistication, sharing 9% of the variance, followed by household income and occupation, which each accounted for 5% of the variance. Household income also accounted for 4% of the variance in lexical diversity, but no other significant SES associations with mothers' speech were observed (Table S6). Sensitivity analyses (see Supporting Information) showed that the effect sizes were robust after controlling for mothers' scores on the reading subtest of the wide-range achievement test (WRAT-3; Snelbaker et al., 2001). Overall, the findings suggest that mothers' SES is more strongly and consistently associated with their vocabulary sophistication than with their lexical diversity and grammar.

Predicting children's cognition and literacy from mothers' SES

Next, we modeled the extent to which mothers' SES predicted children's cognition and literacy (Figure 1B; Tables S2–4). Across ages and measures, mothers' SES indicators accounted for 8%–10% of the variance in children's cognition and literacy (Figure 1B). Mothers' educational attainment and household income accounted independently for 3%–4% of the variance in children's cognition and literacy across ages and measures. Mothers' occupation was not significantly associated with children's cognition and literacy at any age.

Predicting children's educational achievement from mothers' SES

Mothers' educational attainment and household income, respectively, accounted for 9% and 14% of children's stable variance in educational achievement (Table S5) but were not associated with changes in achievement (slope) across the school years. Mothers' occupation was not significantly associated with children's educational achievement. Our findings suggest that mothers' educational attainment and household income can predict children's differences in performance that are already evident at the start of schooling, but not relative changes in achievement.

Discerning predictions from mothers' speech versus SES for children's cognition and literacy

Next, we tested the extent to which our previously observed associations between mothers' vocabulary sophistication and children's cognition and literacy could be attributed to SES (Figure 1; Tables S2–4). Associations between mothers' vocabulary sophistication and children's cognition and literacy remained significant

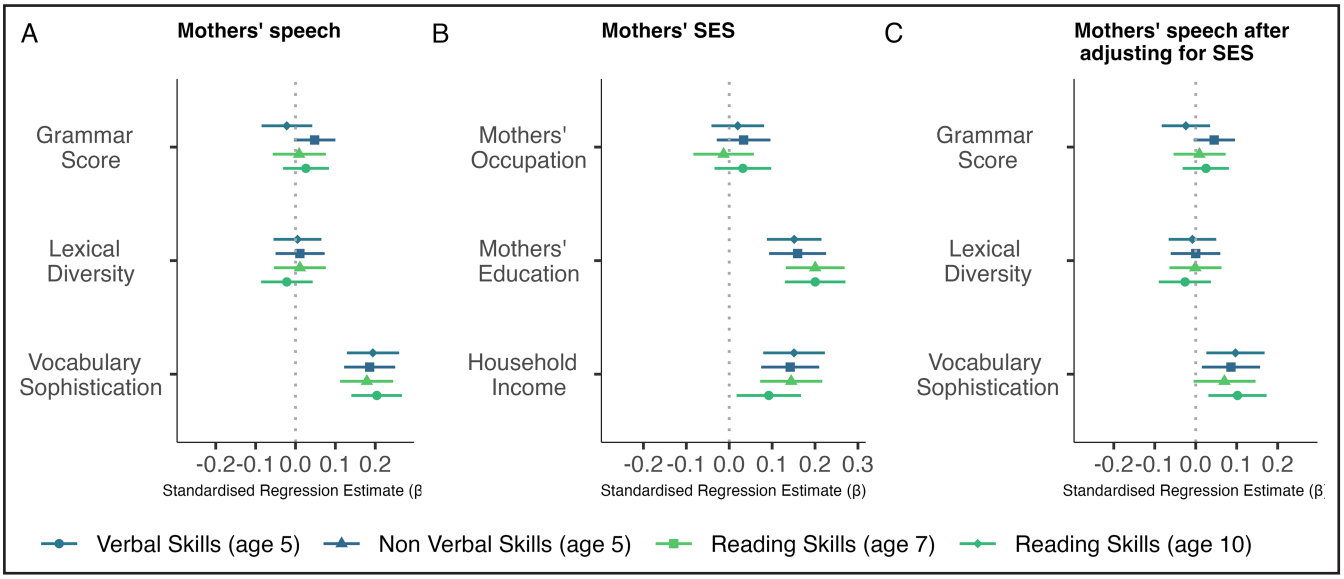

Figure 2 Mothers' education, occupation, and household income as predictors of their speech characteristics. Note: $N = 894$. Dots represent standardized regression (beta) estimates from models where mothers' education, occupation, and income were included as simultaneous predictors after adjusting for mothers' age in years. The lines indicate the 95% confidence intervals.

but halved in effect size (i.e., 1%–2%) after adding mothers' SES to the models. These findings suggest that associations of mothers' SES with children's cognition and literacy are likely more pervasive than those with mothers' vocabulary sophistication. Our sensitivity analyses (see Supporting Information for more details) confirmed that the observed associations were robust and independent of covariate effects. We observed no differences in the associations between mothers' lexical diversity and grammar with children's cognition and literacy, which were nonsignificant before and after adding mothers' SES to the models.

To test whether mothers' vocabulary sophistication mediated associations between SES and children's cognition and literacy (Figure 3A), we specified mothers' vocabulary sophistication as a mediator and mothers' education and household income as predictors (Figure 3A). These SES indicators were significantly associated with children's cognition and literacy in our earlier regression analyses, providing the empirical basis for testing mediation (Baron & Kenny, 1986; MacKinnon et al., 2007; see Supporting Information for full details). Mediation models were used to estimate the indirect effect size of mothers' vocabulary sophistication in the SES–outcome association, estimating the indirect effect directly using bias-corrected bootstrapped confidence intervals (95% CI, 1,000 resamples; Li, 2011), in line with current recommendations (Hayes, 2009; Zhao et al., 2010).

When focusing on the mediation effect ($ß_m$), which refers to the product of the paths that specifies the indirect effect estimate (dotted arrows in Figure 3A). Mothers' vocabulary sophistication significantly mediated the association between mothers' educational attainment and children's verbal ($ß_m^i = 0.04$, 95% CI = 0.01–0.07) and nonverbal skills at age 5 ($ß_m = 0.03$, 95% CI = 0.01–0.05) as well as their literacy skills at age 10 ($ß_m = 0.03$, 95% CI = 0.01–0.06). At age 7, the mediation for literacy skills was not significant ($ß_m = 0.02$, 95% CI = −0.01 to 0.05; Figure 3B and Table S7). Mothers' vocabulary sophistication significantly mediated the association between household income and children's verbal ($ß_m = 0.03$, 95% CI = 0.01–0.05) and nonverbal skills at age 5 ($ß_m = 0.02$, 95% CI = 0.01–0.04), and their literacy skills at age 10 ($ß_m = 0.02$, 95% CI = 0.01–0.05), but not at age 7 ($ß_m = 0.02$, 95% CI = −0.01 to 0.05).

Effect sizes for the significant mediations were small, ranging from 13% to 24% of the associations between SES and children's cognition and literacy, with models accounting for 11% to 8% of the variance (i.e., absolute mediation effect sizes were 3% to 1%; full model results in Table S7). These results suggest that mothers' vocabulary sophistication partly explains the associations between mothers' educational attainment and household income with children's cognition and literacy, albeit with modest effect sizes and inconsistently across measures.

Discerning predictions from mothers' speech versus SES for children's educational achievement

We tested the extent to which associations between mothers' vocabulary sophistication and children's stable variance in educational achievement could be attributed to SES (Table S6). Associations between mothers' vocabulary sophistication and children's educational achievement became nonsignificant after adding mothers' SES to the models, suggesting again that associations between mothers' SES and children's educational achievement are likely more pervasive than those with mothers' vocabulary sophistication.

Because mothers' vocabulary sophistication was independently a significant predictor of children's stable variance in educational achievement, the prerequisite for testing mediation was met. We specified mothers' vocabulary sophistication to mediate the associations between their educational attainment or household income and their children's educational achievement. Mothers' vocabulary sophistication mediated neither association (educational attainment: $ß_m = 0.01$, 95% CI = −0.01 to 0.04; household income: $ß_m = 0.01$, 95% CI = −0.01 to 0.03; Figure 3B; Table S7).

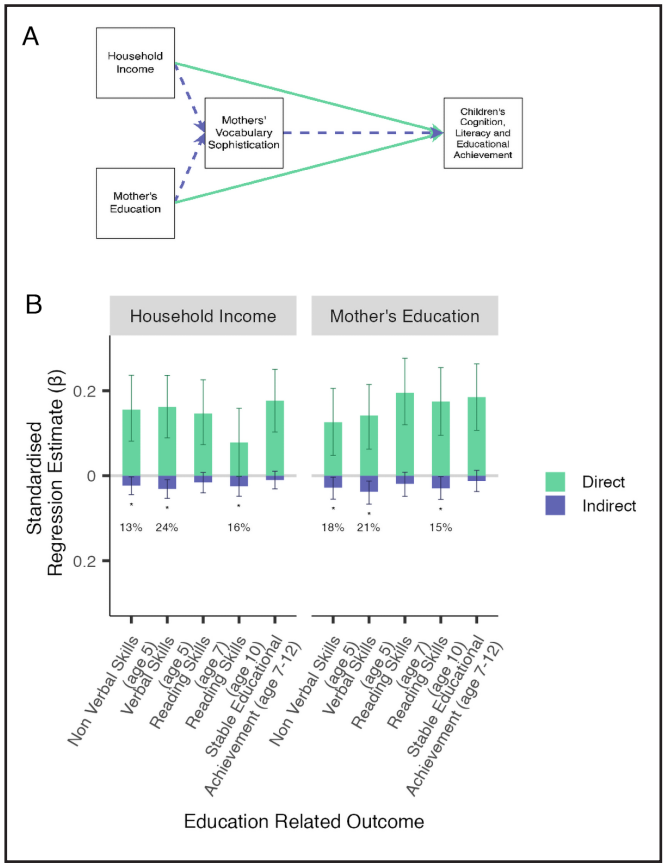

Figure 3 Direct and mediated effects of mothers' vocabulary sophistication on children's cognition, literacy, and educational achievement. Note. $N = 894$, Top panel (A) illustrates the principal mediation model that was fitted. Solidarrows (From SES marker to Children's outcome) represent direct effects; dotted arrows (From SES marker to Children's outcome via mothers' vocabulary sophistication) represent mediated effects. Bottom panel (B) shows the direct estimates above the zero line and mediated estimates below the zero line. For mediated estimates, betas transformed into % of variance explained by the indirect effect, with 95% bootstrapped confidence intervals from 1,000 resamples for mothers' vocabulary sophistication (mediator), separately for mothers' education and household income (predictors) cognition, literacy, and educational achievement (outcomes).

Earlier, we found that mothers' vocabulary sophistication mediated the relation between mothers' SES and children's cognition and literacy. Yet, for educational achievement, the mediations were not significant.

Discussion

Previous research that sought to understand the role of family background for children's educational outcomes studied the language environments that children experienced during their early years (Pace et al., 2017). These studies typically focused on the vocabulary and grammar that characterized the speech that mothers directed at their children before the latter started formal education (e.g., child-directed speech; Huttenlocher et al., 2010; Rowe, 2012; Weizman & Snow, 2001). Our study is the first to examine whether mothers' adult-directed speech, which was assessed when children were 5 years old and about to start primary education, predicts children's educational outcomes across school years. We observed that mothers' vocabulary sophistication, but not two other characteristics of their speech (i.e., lexical diversity and grammatical complexity), predicted children's educational achievement, accounting for 2% of the variance. This small effect attributed to mothers' vocabulary sophistication became nonsignificant after considering mothers' educational attainment and household income (i.e., socioeconomic resources; Oakes & Rossi, 2003). Mothers' socioeconomic resources explained 9% and 14% of children's educational achievement, respectively, a fivefold greater effect than mothers' vocabulary sophistication. These findings may explain why interventions that target mothers' speech to disrupt family background inequality in education often show only modest effects (Adamson et al., 2020).

We contend that mothers' child-directed speech predicts children's language abilities during the early years, as shown in previous studies (d'Apice & von Stumm, 2023; Rowe, 2012), and that these early language abilities aid children's subsequent educational achievement (Durham et al., 2007; Pace et al., 2019). However, findings here and elsewhere (von Stumm et al., 2022a, 2022b) suggest that mothers' speech may exert only a modest direct influence on children's educational achievement. By comparison, mothers' SES is a much stronger predictor of children's educational performance throughout the school years (e.g., Ensminger & Fothergill, 2003; Starr et al., 2024). The narrow focus of past research on the role of mothers' speech for children's educational achievement (Egan-Dailey et al., 2024; Sperry et al., 2019) may have deterred studying alternative linguistic and nonlinguistic pathways that contribute to the transmission of family background inequality in education (Caro et al., 2009). Ignoring these alternative pathways may have reinforced—inadvertently—narratives that mothers, rather than the social and economic conditions in which they raise their children, are at the root of children's educational achievement differences (Adair et al., 2017; Sperry et al., 2019).

Our findings also challenge the idea that mothers' speech differences are shaped by the "social gradient," a widely observed phenomenon in which people with fewer socioeconomic resources are at increased risk of less favorable outcomes (e.g., worse speech; Mulder & Hulstijn, 2011) relative to people with greater resources (Marmot, 2015). We found small associations between mothers' SES and their vocabulary sophistication and lexical diversity, but none with their grammatical complexity. Our findings of a weak social gradient align with recent analyses of other large-scale cohorts (Bergelson et al., 2023) and meta-analytic studies (Anderson et al., 2021; Dailey & Bergelson, 2022) but contradict some reports of dramatic SES-related differences in mothers' speech in smaller samples (Huttenlocher et al., 2010; Rowe, 2008). In our analyses, we found little evidence that vocabulary sophistication, lexical diversity, and grammatical complexity, which are typical markers of adult-directed speech, covaried systematically with education, occupation, and household income. This suggests that previous research may have overestimated the degree to which variability in the structural complexity of mothers' speech can be attributed to their differences in SES.

Our findings elucidate methodological gaps in the current understanding of the role of mothers' speech in the transmission of educational inequality. First, measurement differences between studies may contribute to inconsistencies in how

strongly language markers align with socioeconomic indicators (Fergadiotis & Wright, 2011). Studies may operationalize the same linguistic construct differently through different elicitation contexts, word lists or coding criteria. Even when targeting the same construct, such as vocabulary sophistication, variation in elicitation methods and in the lists or criteria used to define "sophisticated" words may affect observed associations with SES (Mahurin et al., 2021; Mahurin-Smith et al., 2015). It is important to consider how methodological choices shape the extent to which language features capture socioeconomic differences. Second, variability in the features of mothers' speech we examined—vocabulary sophistication, lexical diversity, and grammatical complexity—appears to capture aspects of language use or verbal skill that are not fully explained by SES. Other variability likely arises from factors such as individual differences in verbal ability and cognitive processing (Mulder & Hulstijn, 2011), as well as genetic influences (Dale et al., 2000). Consequently, associations between mothers' and children's speech may arise through both environmental and genetic pathways. In some cases, these genetic pathways may better account for the mother–child association than SES, implying that links between mothers' and children's speech could be genetically confounded rather than environmentally driven. Future research should examine within- and between-family variation in mothers' speech to help clarify the extent to which associations between mothers' and child's speech are genetically entangled, while recognizing that genetic influence does not imply determinism, nor environmental influence straightforward modifiability. Finally, other structural features of mothers' speech not examined here, such as dialect (Curtin, 2020) and narrative construction (Ochs & Schieffelin, 2017), may align with the social gradient. Beyond structural dimensions of language, interactional qualities found in child-directed speech—including maternal sensitivity, contingency, and developmentally tailored scaffolding—are likewise shaped by socioeconomic context *(Madigan* et al., *2019; Rowe & Snow, 2020).* Therefore, a weak SES gradient in mothers' adult-directed speech does not imply a weak SES gradient in child-directed interactions. These interactional qualities may, in turn, mediate or moderate the relationship between structural features of speech and children's outcomes, helping to explain how social gradients in language use translate into differences in educational trajectories.

Our overall finding was that mothers' adult-directed speech likely plays a significant yet modest role in children's cognition, literacy, and educational achievement. Two additional observations emerged. First, among the three characteristics of mothers' speech studied here, vocabulary sophistication (i.e., % of rare words; Weizman & Snow, 2001) was the most consistent predictor of children's educational outcomes. It was also the speech characteristic most strongly associated with mothers' SES. The potential influence of mothers' vocabulary sophistication on their children's vocabulary development has been previously noted (Hoff & Naigles, 2002; Weizman & Snow, 2001). Exposure to sophisticated vocabulary will likely benefit children's learning from academic texts, which become increasingly abstract, complex, and dense across school years (Uccelli, 2023). Thus, our findings support calls for classroom environments that teach sophisticated vocabulary on a par with the academic texts that pupils study (Cabral-Gouveia et al., 2023). That said, the exact pathways and mechanisms through which mothers' vocabulary sophistication may inform children's developmental differences in education-related outcomes are, to date, unknown (Schleppegrell, 2001; Uccelli, 2023).

The other observation was that mothers' vocabulary sophistication predicted children's cognition and literacy more strongly than children's educational achievement. Specifically, mothers' vocabulary sophistication continued to predict children's cognition and literacy after adjusting for mothers' educational attainment and household income, while its prediction of educational achievement became nonsignificant. Also, mothers' vocabulary sophistication partly mediated associations between mothers' educational attainment and household income with children's cognition and literacy, but not with children's educational achievement. Cognition and literacy are specific skills that inform educational achievement (Dickinson et al., 2010; Magalhães et al., 2020), which is a broad concept encompassing various factors, such as school attendance, study habits, and classroom behavior (Ratcliff et al., 2016). Mothers' vocabulary sophistication may have direct benefits for the development of cognition and literacy (Beck et al., 2013; Dickinson et al., 2010), as observed in the current study. By comparison, its influences on educational achievement may be more indirect (Vadivel et al., 2023), and these indirect influences were masked in our analyses.

Limitations

Our study has many strengths, including rich observational data across a 7-year follow-up period from a large sample representative of the full range of Britain's socioeconomic conditions (Nielsen et al., 2017). Yet, it is not without limitations. First, our sample was monolingual and predominantly White British. Because E-Risk does not reflect ethnic minority experiences, our findings should not be generalized to the important intersections between ethnicity, linguistic features, and educational inequality. Second, comparing the lexical and grammatical complexity in mothers' child- versus their adult-directed speech was not possible (c.f. Hoff, 2003; Rowe, 2012), because no measures of mothers' child-directed speech were available. Also, we focused only on mothers, but no data were available on the speech of fathers, teachers, and peers, which may also influence children's educational outcomes (Barnes & Dickinson, 2017; Chen et al., 2020; Pancsofar & Vernon-Feagans, 2006). Thus, our analyses allow inferring conclusions about the role of mothers' adult-directed speech in the transmission of family background inequality in education, but not about other dimensions of children's language experiences, including child-directed interaction, language from other caregivers or teachers, or the overall quantity and quality of linguistic input in the home. Third, although our data are comparatively rich, they were collected infrequently (every 2–5 years), assessing some learning skills (i.e., cognition and literacy) but not others (e.g., self-regulation, numeracy). Moreover, our study's measures of cognition and literacy likely only captured a narrow aspect of their respective broad construct spaces, excluding domains like verbal fluency and reading comprehension. Also, no data on children's cognitive and literacy abilities before the age of 5 years were available, which made it impossible to control for early life language skills that may affect the role of mothers' speech for children's later cognition, literacy, and educational achievement. We therefore caution that, while our findings are clear and consistent, future research needs to replicate and extend them.

Conclusion

Our findings suggest that mothers' education and household income exert a greater long-term influence on children's educational outcomes than the characteristics of mothers' speech. Efforts that seek to disrupt the perpetuation of family background inequality in education should consider a broad range of possible transmission pathways. For example, pedagogical approaches at the classroom level (Cabral-Gouveia et al., 2023; West et al., 2024) can enable all children to maximize the learning opportunities that school affords them, regardless of their cognition and literacy skills at school entry or their families' socioeconomic resources (Adair et al., 2017). A first step to effectively reducing family background inequality in education is recognizing that achievement gaps emerge and manifest in the context of wider, systemic discrimination.

Supplementary material

Supplementary material is available at *Child Development* online.

Data availability

Preregistration: The hypotheses, methods, and analyses were preregistered (https://sites.duke.edu/moffittcaspiprojects/files/2023/05/Brown2023_Early-language-experience-and-inequality_erisk.pdf) prior to data collection. Materials and Data: All materials and data are available upon request to the E-Risk study steering committee (https://eriskstudy.com/data-access/). Analysis scripts: The analysis code used in this study is available on the Open Science Framework (https://osf.io/shcjx).

Author contributions

Anna Brown (Conceptualization, Data curation [supporting], Formal analysis, Writing—original draft [lead], Investigation, Methodology, Writing—review & editing [equal]), Avshalom Caspi (Conceptualization, Data curation, Investigation, Methodology, Writing—review & editing [equal]), Helen L Fisher (Conceptualization, Data curation, Investigation, Methodology, Writing—review & editing [equal]), Terrie Moffitt (Conceptualization, Data curation, Investigation, Methodology, Writing—review & editing [equal]), Jasmin Wertz (Conceptualization, Data curation, Investigation, Methodology, Writing—review & editing [equal]), and Sophie Von Stumm (Conceptualization, Data curation, Investigation, Methodology, Writing—review & editing [equal], Formal analysis, Writing—original draft [supporting])

Funding

This project was funded by the Nuffield Foundation [EDO/FR-000023238]. The Environmental Risk (E-Risk) Study is funded by grants from the Medical Research Council (MRC) [G1002190; MR/X010791/1]. Additional support was provided by the National Institute of Child Health and Human Development [HD077482] and the Jacobs Foundation. Sophie von Stumm was a Jacobs Foundation CRISP Fellow (2022–2027) and a Paris Institute of Advanced Study Fellow (2023–2024) during the writing of this manuscript. Helen L. Fisher was part supported by the Economic and Social Research Council (ESRC) Centre for Society and Mental Health at King's College London [ES/S012567/1; UKRI861].

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgments

We are grateful to the study mothers and fathers, the twins, and the twins' teachers for their participation. Our thanks to Professor Louise Arseneault, a former principal investigator of the E-Risk study, and to the E-Risk team for their dedication, hard work, and insights.

References

Adair, J. K., Colegrove, K. S.-S., & McManus, M. E. (2017). How the word gap argument negatively impacts young children of Latinx immigrants' conceptualizations of learning. *Harvard Educational Review*, *87*, 309–334. https://doi.org/10.17763/1943-5045-87.3.309

Adamson, L. B., Kaiser, A. P., Tamis-LaMonda, C. S., Owen, M. T., & Dimitrova, N. (2020). The developmental landscape of early parent-focused language intervention. *Early Childhood Research Quarterly*, *50*, 59–67. https://doi.org/10.1016/j.ecresq.2018.11.005

Akhtar, N. (2005). The robustness of learning through overhearing. *Developmental Science*, *8*, 199–209. https://doi.org/10.1111/j.1467-7687.2005.00406.x

Anderson, N. J., Graham, S. A., Prime, H., Jenkins, J. M., & Madigan, S. (2021). Linking quality and quantity of parental linguistic input to child language skills: A meta-analysis. *Child Development*, *92*, 484–501. https://doi.org/10.1111/cdev.13508

Babayiğit, S. (2015). The relations between word reading, oral language, and reading comprehension in children who speak English as a first (L1) and second language (L2): a multigroup structural analysis. *Reading and Writing*, *28*, 527–544. https://doi.org/10.1007/s11145-014-9536-x

Barnes, E. M., & Dickinson, D. K. (2017). The relationship of head start teachers' academic language use and children's receptive vocabulary. *Early Education and Development*, *28*, 794–809. https://doi.org/10.1080/10409289.2017.1340069

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction*. Guilford Press.

Bergelson, E., Soderstrom, M., Schwarz, I.-C., Rowland, C. F., Ramírez-Esparza, N., R. Hamrick, L., Marklund, E., Kalashnikova, M., Guez, A., Casillas, M., Benetti, L., van Alphen, P., & Cristia, A. (2023). Everyday language input and production in 1,001 children from six continents. *Proceedings of the National Academy of Sciences*, *120*,e2300671120. https://doi.org/10.1073/pnas.2300671120

Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258). Greenwood.

Cabral-Gouveia, C., Menezes, I., & Neves, T. (2023). Educational strategies to reduce the achievement gap: A systematic review. *Frontiers in Education*, *8*, Article 1155741. https://doi.org/10.3389/feduc.2023.1155741

Carlson, R., & McChesney, C. (2015). *Income sustainability through educational attainment* (SSRN Scholarly Paper No. 2676859). The Exchange, Vol. 4, No. 1. https://papers.ssrn.com/abstract=2676859

Caro, D. H., McDonald, J. T., & Willms, J. D. (2009). Socio-economic status and academic achievement trajectories from childhood to adolescence. *Canadian Journal of Education/Revue Canadienne de l'éducation*, *32*, 558–590. https://cje-rce.ca/index.php/cje-rce/article/view/3052

Caspi, A., Moffitt, T. E., Morgan, J., Rutter, M., Taylor, A., Arseneault, L., Tully, L., Jacobs, C., Kim-Cohen, J., & Polo-Tomas, M. (2004). Maternal expressed emotion predicts children's antisocial behavior problems: Using monozygotic-twin differences to identify environmental effects on behavioral development. *Developmental Psychology*, *40*, 149–161. https://doi.org/10.1037/0012-1649.40.2.149

Chall, J. S., & Dale, E. (1995). *Readability revisited: The new Dale-Chall readability formula*. Brookline Books. https://cir.nii.ac.jp/crid/1130282268845043712

Chen, J., Justice, L. M., Tambyraja, S. R., & Sawyer, B. (2020). Exploring the mechanism through which peer effects operate in preschool classrooms to influence language growth. *Early Childhood Research Quarterly*, *53*, 1–10. https://doi.org/10.1016/j.ecresq.2020.02.002

Curtin, M. (2020). Dialect: Social class. *The International Encyclopedia of Linguistic Anthropology*, 1–9. https://doi.org/10.1002/9781118786093.iela0100

Dailey, S., & Bergelson, E. (2022). Language input to infants of different socioeconomic statuses: A quantitative meta-analysis. *Developmental Science*, *25*,e13192. https://doi.org/10.1111/desc.13192

Dale, P. S., Dionne, G., Eley, T. C., & Plomin, R. (2000). Lexical and grammatical development: A behavioural genetic perspective. *Journal of Child Language*, *27*, 619–642. https://doi.org/10.1017/S0305000900004281

d'Apice, K., & von Stumm, S. (2023). Does age moderate the influence of early life language experiences? A naturalistic home observation study. *Early Childhood Research Quarterly*, *63*, 400–409. https://doi.org/10.1016/j.ecresq.2023.01.009

Dickinson, D. K., Golinkoff, R. M., & Hirsh-Pasek, K. (2010). Speaking out for language: Why language is central to reading development. *Educational Researcher*, *39*, 305–310. https://doi.org/10.3102/0013189X10370204

Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry*, *56*, 848–856. https://doi.org/10.1111/jcpp.12378

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*, 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

Durham, R. E., Farkas, G., Hammer, C. S., Bruce Tomblin, J., & Catts, H. W. (2007). Kindergarten oral language skill: A key variable in the intergenerational transmission of socioeconomic status. *Research in Social Stratification and Mobility*, *25*, 294–305. https://doi.org/10.1016/j.rssm.2007.03.001

Egan-Dailey, S., Gennetian, L. A., Magnuson, K., Duncan, G. J., Yoshikawa, H., Fox, N. A., & Noble, K. G. (2024). Child-directed speech in a large sample of U.S. mothers with low income. *Child Development*, *95*, 2045–2061. https://doi.org/10.1111/cdev.14139

Ensminger, M., & Fothergill, K. E. (2003). A decade of measuring S.E.S: What it tells us and where to go from here. In M. H. Bornstein, & R. H. Bradley (Eds.), *Socioeconomic status parenting, and child development* (pp. 13–27). Lawrence Erlbaum Associates.

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2005). First grade and educational attainment by age 22: A new story. *The American Journal of Sociology*, *110*, 1458–1502. https://doi.org/10.1086/428444

Fang, Z., Schleppegrell, M. J., & Cox, B. E. (2006). Understanding the language demands of schooling: Nouns in academic registers. *Journal of Literacy Research*, *38*, 247–273. https://doi.org/10.1207/s15548430jlr3803_1

Fergadiotis, G., & Wright, H. H. (2011). Lexical diversity for adults with and without aphasia across discourse elicitation tasks. *Aphasiology*, *25*, 1414. https://doi.org/10.1080/02687038.2011.603898

Furnée, C. A., Groot, W., & van den Brink, H. M. (2008). The health effects of education: A meta-analysis. *European Journal of Public Health*, *18*, 417–421. https://doi.org/10.1093/eurpub/ckn028

Girouard-Hallam, L. N., & Norris, M. N. (2024). Overheard and understood: A systematic review of children's learning from overhearing. *Translational Issues in Psychological Science*, *12*, 129–142. https://doi.org/10.1037/tps0000399.

Golden, C. J., & Bennett, R. (2024). *Clinical integration of neuropsychological test results*. CRC Press.

Great Britain; Office of Population Censuses, Surveys, & Great Britain; Government Statistical Service. (1995). *Standard occupational classification* (Vol. *2*). HM Stationery Office.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, *60*, 549–576. https://doi.org/10.1146/annurev.psych.58.110405.085530

Gustavson, K., von Soest, T., Karevold, E., & Røysamb, E. (2012). Attrition and generalizability in longitudinal studies: Findings from a 15-year population-based study and a Monte Carlo simulation study. *BMC Public Health*, *12*,918. https://doi.org/10.1186/1471-2458-12-918

Hadley, E. B., Barnes, E. M., Wiernik, B. M., & Raghavan, M. (2022). A meta-analysis of teacher language practices in early childhood classrooms. *Early Childhood Research Quarterly*, *59*, 186–202. https://doi.org/10.1016/j.ecresq.2021.12.002

Halliday, M. A. K., & Hasan, R. (1989). *Language, context and text: Aspects of language in a social-semiotic perspective* (2nd ed.). Oxford University Press.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. P.H. Brookes.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, *76*, 408–420. https://doi.org/10.1080/03637750903310360

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development

via maternal speech. *Child Development*, *74*, 1368–1378. https://doi.org/10.1111/1467-8624.00612

Hoff, E. (2013). Interpreting the early language trajectories of children from low SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology*, *49*, 4–14. https://doi.org/10.1037/a0027238

Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development*, *73*, 418–433. https://doi.org/10.1111/1467-8624.00415

Hoff-Ginsberg, E. (1990). Maternal speech and the child's development of syntax: A further look. *Journal of Child Language*, *17*, 85–99. https://doi.org/10.1017/S0305000900013118

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*, 343–365. https://doi.org/10.1016/j.cogpsych.2010.08.002

Killingly, C., Graham, L. J., Tancredi, H., & Snow, P. (2025). Reciprocal relationships among reading and vocabulary over time: a longitudinal study from grade 1 to 5. *Reading and Writing*, *38*, 605–625. https://doi.org/10.1007/s11145-024-10522-x

Kircher, R., & Fox, S. (2021). Multicultural London English and its speakers: A corpus-informed discourse study of standard language ideology and social stereotypes. *Journal of Multilingual and Multicultural Development*, *42*, 792–810. https://doi.org/10.1080/01434632.2019.1666856

Lareau, A., & Weininger, E. B. (2003). Cultural capital in educational research: A critical assessment. *Theory and Society*, *32*, 567–606. https://doi.org/10.1023/B:RYSO.0000004951.04408.b0

Li, S. D. (2011). Testing mediation using multiple regression and structural equation modeling analyses in secondary data. *35*, 240–268. https://doi.org/10.1177/0193841X11412069

Lippi-Green, R. (2011). *English with an accent: Language, ideology and discrimination in the United States* (2nd ed.). Routledge. https://doi.org/10.4324/9780203348802

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, *58*, 593. https://doi.org/10.1146/annurev.psych.58.110405.085542

MacWhinney, B. (2000). The CHILDES project. *Computational Linguistics*, *26*, 657–657. https://doi.org/10.1162/coli.2000.26.4.657

Madigan, S., Prime, H., Graham, S. A., Rodrigues, M., Anderson, N., Khoury, J., & Jenkins, J. M. (2019). Parenting behavior and child language: A meta-analysis. *Pediatrics*, *144*,e20183556. https://doi.org/10.1542/peds.2018-3556

Magalhães, S., Carneiro, L., Limpo, T., & Filipe, M. (2020). Executive functions predict literacy and mathematics achievements: The unique contribution of cognitive flexibility in grades 2, 4, and 6. *Child Neuropsychology*, *26*, 934–952. https://doi.org/10.1080/09297049.2020.1740188

Magnuson, K., Duncan, G. J., Lee, K. T. H., & Metzger, M. W. (2016). Early school adjustment and educational attainment. *American Educational Research Journal*, *53*, 1198–1228. https://doi.org/10.3102/0002831216634658

Mahurin Smith, J., DeThorne, L., & Petrill, S. (2015). From aardvark to ziggurat: A new tool for assessing children's use of rare vocabulary. *Clinical Linguistics & Phonetics*, *29*, 441–454. https://doi.org/10.3109/02699206.2015.1016187

Mahurin-Smith, J., Mills, M. T., & Chang, R. (2021). Rare vocabulary production in school-age narrators from low-income communities. *Language, Speech, and Hearing Services in Schools*, *52*, 51–63. https://doi.org/10.1044/2020_LSHSS-19-00120

Marmot, M. (2015). The health gap: The challenge of an unequal world. *The Lancet*, *386*, 2442–2444. https://doi.org/10.1016/S0140-6736(15)00150-6

McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, *42*, 381–392. https://doi.org/10.3758/BRM.42.2.381

McKee, G., Malvern, D., & Richards, B. (2000). Measuring vocabulary diversity using dedicated software. *Literary and Linguistic Computing*, *15*, 323–338. https://doi.org/10.1093/llc/15.3.323

Mercer, N., Wegerif, R., & Dawes, L. (1999). Children's talk and the development of reasoning in the classroom. *British Educational Research Journal*, *25*, 95–111. https://doi.org/10.1080/0141192990250107

Moffitt, T. E.; E-Risk Study Team. (2002). Teen-aged mothers in contemporary Britain. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *43*, 727–742. https://doi.org/10.1111/1469-7610.00082

Mueller, K. D., Koscik, R. L., Clark, L. R., Hermann, B. P., Johnson, S. C., & Turkstra, L. S. (2018). The latent structure and test–retest stability of connected language measures in the Wisconsin Registry for Alzheimer's Prevention (WRAP). *Archives of Clinical Neuropsychology*, *33*, 993–1005. https://doi.org/10.1093/arclin/acx116

Mulder, K., & Hulstijn, J. H. (2011). Linguistic skills of adult native speakers, as a function of age and level of education. *Applied Linguistics*, *32*, 475–494. https://doi.org/10.1093/applin/amr016

Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology*, *162*, 31–38. https://doi.org/10.1016/j.jecp.2017.04.017

Noble, C., Medin, D., Quail, Z., Young, C., & Carter, M. (2021). How does participation in formal education or learning for older people affect wellbeing and cognition? A systematic literature review and meta-analysis. *Gerontology and Geriatric Medicine*, *7*,2333721420986027. https://doi.org/10.1177/2333721420986027

Oakes, J. M., & Rossi, P. H. (2003). The measurement of SES in health research: Current practice and steps toward a new approach. *Social Science & Medicine (1982)*, *56*, 769–784. https://doi.org/10.1016/s0277-9536(02)00073-4

Ochs, E., & Schieffelin, B. (2017). Language socialization: An historical overview. In P. A. Duff, & S. May (Eds.), *Language socialization* (pp. 1–14). Springer International Publishing. https://doi.org/10.1007/978-3-319-02327-4_1-2

Odgers, C. L., Caspi, A., Russell, M. A., Sampson, R. J., Arseneault, L., & Moffitt, T. E. (2012). Supportive parenting mediates neighborhood socioeconomic disparities in children's antisocial behavior from ages 5 to 12. *Development and Psychopathology*, *24*, 705–721. https://doi.org/10.1017/S0954579412000326

Pace, A., Alper, R., Burchinal, M. R., Golinkoff, R. M., & Hirsh-Pasek, K. (2019). Measuring success: Within and cross-domain predictors of academic and social trajectories in elementary school. *Early Childhood Research Quarterly*, *46*, 112–125. https://doi.org/10.1016/j.ecresq.2018.04.001

Pace, A., Luo, R., Hirsh-Pasek, K., & Golinkoff, R. M. (2017). Identifying pathways between socioeconomic status and language development. *Annual Review of Linguistics*, *3*, 285–308. https://doi.org/10.1146/annurev-linguistics-011516-034226

Pancsofar, N., & Vernon-Feagans, L. (2006). Mother and father language input to young children: Contributions to later language development. *Journal of Applied Developmental Psychology*, *27*, 571–587. https://doi.org/10.1016/j.appdev.2006.08.003

Preacher, K. J. (2018). Latent growth curve models. In *The reviewer's Guide to quantitative methods in the social sciences* (pp. 178–192). Routledge.

Ratcliff, N. J., Costner, R. H., Carroll, K. L., Jones, C. R., Sheehan, H. C., & Hunt, G. H. (2016). Causes of and solutions to the achievement gap: Teachers' perceptions. *Teacher Educators Journal*, *9*, 97–111.

Rimfeld, K., Malanchini, M., Hannigan, L. J., Dale, P. S., Allen, R., Hart, S. A., & Plomin, R. (2019). Teacher assessments during compulsory education are as reliable, stable and heritable as standardized test scores. *Journal of Child Psychology and Psychiatry*, *60*, 1278–1288. https://doi.org/10.1111/jcpp.v60.12

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1–36. https://doi.org/10.18637/jss.v048.i02

Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill*. *Journal of Child Language*, *35*, 185–205. https://doi.org/10.1017/S0305000907008343

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, *83*, 1762–1774. https://doi.org/10.1111/j.1467-8624.2012.01805.x

Rowe, M. L. (2018). Understanding socioeconomic differences in parents' speech to children. *Child Development Perspectives*, *12*, 122–127. https://doi.org/10.1111/cdep.12271

Rowe, M. L., & Snow, C. E. (2020). Analyzing input quality along three dimensions: Interactive, linguistic, and conceptual. *Journal of Child Language*, *47*, 5–21. https://doi.org/10.1017/S0305000919000655

Rutter, M., Caspi, A., Fergusson, D., Horwood, L. J., Goodman, R., Maughan, B., Moffitt, T. E., Meltzer, H., & Carroll, J. (2004). Sex Differences in Developmental Reading Disability. *JAMA*, *291*, 2007. https://doi.org/10.1001/jama.291.16.2007

Schleppegrell, M. J. (2001). Linguistic features of the language of schooling. *Linguistics and Education*, *12*, 431–459. https://doi.org/10.1016/S0898-5898(01)00073-0

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Routledge. https://www.taylorfrancis.com/books/mono/10.4324/9781410610317/language-schooling-mary-schleppegrell

Schwab, J. F., & Lew-Williams, C. (2016). Language learning, socioeconomic status, and child-directed speech. *Wiley Interdisciplinary Reviews: Cognitive Science*, *7*, 264–275. https://doi.org/10.1002/wcs.1393

Snelbaker, A. J., Wilkinson, G. S., Robertson, G. J., & Glutting, J. J. (2001). Wide range achievement test 3 (wrat3). In W. I. Dorfman, & M. Hersen (Eds.), *Understanding psychological assessment* (pp. 259–274). Springer US. https://doi.org/10.1007/978-1-4615-1185-4_13

Spencer, S., Clegg, J., & Stackhouse, J. (2013). Language, social class and education: Listening to adolescents' perceptions. *Language and Education*, *27*, 129–143. https://doi.org/10.1080/09500782.2012.760585

Sperry, D. E., Sperry, L. L., & Miller, P. J. (2019). Reexamining the verbal environments of children from different socioeconomic backgrounds. *Child Development*, *90*, 1303–1318. https://doi.org/10.1111/cdev.13072

Starr, A., Haider, Z. F., & von Stumm, S. (2024). Do school grades matter for growing up? Testing the predictive validity of school performance for outcomes in emerging adulthood. *Developmental Psychology*, *60*, 665–679. https://doi.org/10.1037/dev0001548

Tarar, J. M, Meisinger, E. B., & Dickens, R. H. (2015). Test Review: *Test of Word Reading Efficiency–Second Edition* (TOWRE-2) by Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. *Canadian Journal of School Psychology*, *30*, 320–326. https://doi.org/10.1177/0829573515594334

Torgesen, J. K., Rashotte, C. A., & Wagner, R. K. (1999). *TOWRE: Test of word Reading efficiency*. PRO-ED.

Trouton, A., Spinath, F. M., & Plomin, R. (2002). Twins early development study (TEDS): A multivariate, longitudinal genetic investigation of language, cognition and behavior problems in childhood. *Twin Research: The Official Journal of the International Society for Twin Studies*, *5*, 444–448. https://doi.org/10.1375/136905202320906255

Uccelli, P. (2023). The language demands of analytical reading and writing at school. *Written Communication*, *40*, 518–554. https://doi.org/10.1177/07410883221148727

Vadivel, B., Alam, S., Nikpoo, I., & Ajanil, B. (2023). The impact of low socioeconomic background on a child's educational achievements. *Education Research International*, *2023*,6565088. https://doi.org/10.1155/2023/6565088

Verhoeven, L., van Leeuwe, J., & Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading*, *15*, 8–25. https://doi.org/10.1080/10888438.2011.536125

von Stumm, S., Cave, S. N., & Wakeling, P. (2022a). Persistent association between family socioeconomic status and primary school performance in Britain over 95 years. *NPJ Science of Learning*, *7*, 4. https://doi.org/10.1038/s41539-022-00120-3

von Stumm, S., O'Reilly, J., & d'Apice, K. (2022b). Predicting developmental outcomes in middle childhood from early life language and parenting experiences. *The British Journal of Developmental Psychology*, *40*, 487–503. https://doi.org/10.1111/bjdp.12427

Wechsler, D. (1989). *Manual for the Wechsler preschool and primary scale of intelligence—revised*. The Psychological Corporation.

Weizman, Z. O., & Snow, C. E. (2001). Lexical output as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology*, *37*, 265–279. https://doi.org/10.1037/0012-1649.37.2.265

West, G., Lervåg, A., Birchenough, J. M. H., Korell, C., Rios Diaz, M., Duta, M., Cripps, D., Gardner, R., Fairhurst, C., & Hulme, C. (2024). Oral language enrichment in preschool improves children's language skills: A cluster randomised controlled trial. *Journal of Child Psychology and Psychiatry*, *65*, 1087–1097. https://doi.org/10.1111/jcpp.13947

Zhao, X., Lynch, J. G. Jr, & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *The Journal of Consumer Research*, *37*, 197–206. https://doi.org/10.1086/651257